

UNIVERSIDADE DE LISBOA

FACULDADE DE LETRAS



# **O TROPISMO AUTONÓMICO PARA A VERDADE**

**JOÃO FRANCISCO DOMINGOS PINHEIRO**

Tese orientada pelo Prof. Doutor António José Teiga Zilhão, especialmente elaborada para a obtenção do grau de Mestre em Filosofia.

2018

## Resumo

Contra o anti-realismo, uma explicação comum do sucesso predictivo da Ciência diz-nos que as teorias e modelos científicos são, pelo menos, aproximadamente verdadeiras acerca do Mundo. Todavia, a natureza exacta da relação geral entre sucesso predictivo e verdade das representações continua por explorar. Reconhecendo que esta é uma relação aberta à investigação empírica, neste ensaio damos os primeiros passos para a construção de uma teoria naturalista do sucesso predictivo da Ciência. Começamos por entender o sucesso predictivo de representações como um fenómeno natural, recorrendo à teoria autonómica e à teoria evolutiva para estudá-lo. É a partir da análise permitida por estas teorias que formulamos a tese do tropismo autonómico para a verdade. A tese do tropismo autonómico para a verdade segue-se da natureza autonómica dos biosistemas, resultando da compreensão do modo como estes geram representações sob condições de sucesso interactivo. Com base nesta tese, demonstramos o credo pragmatista de que representações verdadeiras são as únicas a conferir sucesso sistemático às acções que se guiam por representações. Da mesma tese derivamos depois um corolário de relevo para a teoria evolutiva: de que só há selecção causal de mecanismos cognitivos se estes contiverem tropismos para a verdade. Se este corolário é verdadeiro, então ele permite explicar porque é que os nossos mecanismos cognitivos evoluíram de tal forma que produzem teorias e modelos científicos aproximadamente verdadeiros acerca do Mundo, ainda que a exacta história selectiva destes mecanismos continue um mistério. Uma vez que a nossa tese providencia uma explicação naturalista compatível com o Realismo científico, para o sucesso predictivo da Ciência, antecipamos que ela possa constituir o fundamento de uma teoria naturalista e realista do seu sucesso.

Palavras-chave: Sucesso predictivo da Ciência; Tropismo Autonómico para a Verdade; Evolução cognitiva.

## Abstract

Against anti-realism, a common explanation of Science's predictive success tells us that our scientific theories and models are at least approximately true about the World. However, the exact nature of the general relation between predictive success and the representations' truth is underexplored. Recognizing that this relation is open to empirical research, in this essay we give the first steps towards the construction of a naturalistic theory of the predictive success of Science. We start by understanding the predictive success of representations as a natural phenomenon, using the theory of biological autonomy and evolutionary theory to study it. It is from within the analysis made possible by these theories that we go on to formulate the thesis of the autonomic tropism towards truth. The thesis of the autonomic tropism towards truth follows from the autonomic nature of biosystems, resulting from the understanding of how biosystems generate representations under conditions of interactive success. Having this thesis as our foundations, we prove the pragmatist creed that only true representations can confer systematic

success to actions guided by representations. From that same thesis we go on to derive a corollary with relevance to evolutionary theory: there is only causal selection of cognitive mechanisms if they are truth-tropic. If this corollary is true, then we can explain why it is that our cognitive mechanisms evolved in such a way as to produce scientific theories and models which are at least approximately true of the World, although the exact selective history of these mechanisms remains a mystery. Because our thesis provides us with a naturalistic explanation compatible with Scientific realism for the predictive success of Science, we anticipate that it may come to constitute the foundations for a naturalist and realist theory of the success of Science.

Keywords: Predictive success of Science; Autonomic tropism towards truth; Cognitive evolution.

Em memória das minhas avós, Conceição e Palmira.

# Índice

Resumo/Abstract .....	1
Índice .....	4
Agradecimentos.....	6
Preâmbulo (e Proémio) .....	6
<b>I. INTRODUÇÃO À TEORIA INTERACTIVISTA DOS BIOSISTEMAS E DA COGNIÇÃO .....</b>	<b>8</b>
1. A Natureza Interactiva dos Biosistemas .....	8
2. A Natureza Evolutiva dos Biosistemas.....	11
2.1. Evolução: A Dimensão Histórica da Autonomia.....	11
2.2. Autonomia e Fitness .....	13
3. A Teoria Interactivista e Evolucionária das Funções Biológicas .....	21
4. A Natureza e Função Mínima da Cognição .....	26
<b>II. FORMULAÇÃO DA TESE DO TROPISMO AUTONÓMICO PARA A VERDADE .....</b>	<b>30</b>
1. Antecipações como Portadores Primitivos de Conteúdo Pré-cognitivo.....	30
2. Antecipações como Veículos de Representação Desacoplados do Ambiente .....	34
3. A Tese do Tropismo Autonómico para a Verdade .....	37
<b>III. A APLICAÇÃO DA TESE DO TROPISMO AUTONÓMICO PARA A VERDADE À SEMÂNTICA .....</b>	<b>38</b>
1. Uma Semântica de Sucesso ou Pragmatista .....	38
2. O Princípio de Ramsey nas Fundações da Semântica de Sucesso .....	38
3. Uma Semântica de Sucesso Interactivo: o Princípio de Ramsey Naturalizado.....	42
4. A Natureza Epistémica e Ontológica da Verdade.....	43
5. A Natureza dos Conteúdos de Volições e a Subordinação da sua Funcionalidade à Autonomia .....	46
6. Contraexemplos à Semântica de Sucesso Interactivista e Respostas .....	48
<b>IV. APLICAÇÃO DA TESE DO TROPISMO AUTONÓMICO PARA A VERDADE À EPISTEMOLOGIA... 56</b>	<b>56</b>
1. O Corolário Evolucionário do Tropismo Autonómico para a Verdade: a Ponte Mílvia.....	56
2. Argumentos Evolucionários a Favor e Contra a Fiabilidade da Cognição.....	56
3. A Distinção entre “Seleccção <i>Simpliciter</i> ” e “Seleccção Causal” .....	64
4. A Complementaridade das Explicações com Base na Fitness e com Base em Tropismos para a Verdade.....	66
5. Constrangimentos Lógicos, Económicos e Ecológicos à Seleccção Natural de Mecanismos Cognitivos com Tropismos para a Verdade.....	71
6. Contraexemplos à Ponte Mílvia e Respostas .....	77
<b>V. PROLEGÓMENOS A UMA TEORIA NATURALISTA DO SUCESSO DA CIÊNCIA .....</b>	<b>85</b>
Anexo A – Como a Evolução Previne o Surgimento de Demónios Darwinianos: Mecanismos Não-Selectivos .....	94

Anexo B – Como a Evolução Previne o Surgimento de Demónios Darwinianos: Selecção Não-Adaptacionista .....	95
Referências Bibliográficas .....	96

## Agradecimentos

A António Zilhão pela orientação e correcções linguísticas; a Alberto Cordero, David Papi-neau, Elliott Sober, Gil Santos, João Cordovil, Nathalie Gontier, e Rui Diogo, pelas discussões mais extensas; a muitos mais (que terão de ficar por nomear) por diversas razões; e por me darem a oportunidade de estudar a seu custo próprio, agradeço por fim aos meus pais.

## Preâmbulo (e Proémio)

A investigação aqui apresentada partiu da observação da capacidade do *H. sapiens* produzir Ciência e de aplicá-la enquanto tecnologia, intervindo no mundo e *prevendo* o futuro [cf. Staley 2014:171]. Estamos em acordo com uma maioria para quem o *sucesso da Ciência*, que não conhece precedentes terrestres, permanece um dos aspectos mais intrigantes da evolução humana. Este mistério resulta da pergunta: como é que os humanos desenvolveram evolutivamente a sua capacidade de produzir Ciência, que lhes permitiu explorar mundos previamente inacessíveis relativamente aos quais não parecem estar *adaptados* [Delbrück 1986:280]? Ou então, o que é ainda mais intrigante: se não evoluímos para produzir Ciência – conforme muitos autores argumentam – porque é que deveríamos *confiar* nestas nossas capacidades e como poderemos explicar o seu sucesso?

É tendo estas “perguntas horrendas” (*horrid*) por pano de fundo (a expressão é de Darwin [1881a]) que se impõe a tarefa a que dedicámos este ensaio: avançar na construção de alguns fundamentos para uma teoria *naturalista* do sucesso da Ciência, entendendo por naturalista a *posição* ou credo de que a Ciência tem algo a dizer acerca do seu próprio sucesso e que a análise do sucesso da Ciência não deve ser deixada *unicamente* para a Filosofia<sup>1</sup>. Esta tarefa foi bem delimitada por Hilary Kornblith:

«[A]gora temos de tentar explicar como é que criaturas com as faculdades cognitivas que a ciência diz que temos vieram a compreender o tipo de mundo que as ciências geralmente dizem que habitamos.» [1994:43]<sup>2</sup>

Neste contexto é necessário ter em mente que existem várias explicações para o sucesso da Ciência. Talvez a mais famosa seja a de que as teorias e modelos científicos têm sucesso porque são pelo menos *parcialmente* ou *aproximadamente* verdadeiras acerca do Mundo. Segundo Hilary Putnam [1975:73], podemos inferir esta como *a melhor explicação* do sucesso da Ciência porque, contrariamente às restantes, ela não apresenta o sucesso da Ciência como um milagre<sup>3</sup>.

Porém, mesmo que aceitemos esta *abdução* ou *inferência para a melhor explicação* [cf. Lipton 1991 e Zilhão 2010:§1.3.-1.5.1.] de Putnam como correcta, não estaremos ainda na posse de uma explicação naturalista do sucesso da Ciência. De facto, temos ainda que dar conta da justificação *empírica* [cf. Laudan 1981] pela qual as nossas teorias ou modelos serem, pelo menos, aproximadamente verdadeiros acerca do mundo é conducente ao seu sucesso. Felizmente,

---

<sup>1</sup> Para os nossos presentes propósitos entenderemos o Naturalismo como a *posição* definida em corpo de texto; todavia, cf.: e.g. Worrall 1999, Giere 2006, Maddy 2007, Sklar 2010, e Cordero 2013 & 2015.

<sup>2</sup> «[W]e must now try to explain how creatures with the faculties cognitive science tells us we have could have come to understand the kind of world which the sciences generally tell us that we inhabit».

<sup>3</sup> Este é também o principal argumento a favor do Realismo científico, cf. van Fraassen 1980, Psillos 1999, Ladyman et al. 2007, e Chakravartty 2007:10-1 para desenvolvimentos.

Putnam também formulou de um modo genérico a pergunta que poderá servir de fio condutor ao estudo científico da relação entre sucesso e verdade:

«como é que o facto das nossas crenças serem (aproximadamente) verdadeiras explica a nossa sobrevivência?»<sup>4</sup> [1981:39]

Com o intuito de responder a esta pergunta combinámos análises efectuadas no âmbito da *Teoria Autonómica* e da *Teoria Evolutiva*. A escolha destas duas teorias deve-se ao facto de ambas procurarem explicar, entre outras coisas, o sucesso natural dos biosistemas, sendo aplicáveis ao estudo dos mecanismos cognitivos (e sua contribuição para o sucesso biológico dos seus portadores). Sendo os nossos mecanismos cognitivos também responsáveis pela nossa produção da Ciência, o recurso a estas duas teorias surge naturalmente<sup>5</sup>.



Estas reflexões e perguntas constituem o mote da investigação cujo produto aqui apresentamos. Para nosso desagrado, a nossa investigação está inacabada, de tal forma que não reunimos condições para apresentar uma teoria naturalista do sucesso da Ciência. A edificação desta teoria continuará na nossa agenda de investigação; entretanto, apresentamos aqui o resultado da investigação levada a cabo tendo por fio condutor a pergunta de Putnam. Assim, expomos unicamente alguns dos fundamentos que, esperançosamente, poderão constituir a rampa de lançamento para uma teoria naturalista do sucesso da Ciência.

---

<sup>4</sup> «[H]ow does the fact that our beliefs are (approximately) true explain our survival?»

<sup>5</sup> Para adoptar esta pergunta ao estudo do sucesso da Ciência bastará tomar as teorias e os modelos como sistemas de crenças.



# I. INTRODUÇÃO À TEORIA INTERACTIVISTA DOS BIOSISTEMAS E DA COGNIÇÃO

O estudo da natureza físico-química dos sistemas biológicos tomou um sentido próprio com as primeiras teorias biofísicas [e.g. Lotka 1922a e 1992b, e Schrödinger 1944], com os primeiros autores ciberneticistas [e.g. Ashby 1947], e ainda um novo sentido aquando da formação da Teoria dos Sistemas [e.g. von Bertalanffy 1968]. Divergindo em bastantes aspectos, todas estas teorias parecem, não obstante, convergir em alguns pontos-chave, de que são exemplo a ênfase que colocam na análise do comportamento *termodinâmico* dos biosistemas, do *controlo*, e do comportamento *sistémico*. Longe de procurar sintetizar estas diferentes teorias e outras que delas advieram, ou mesmo de procurar acompanhar as mútuas dependências do seu desenvolvimento histórico, o que nos propomos aqui investigar são alguns dos traços substanciais comuns a elas que lhes conferiram aplicabilidade ao estudo dos biosistemas e, posteriormente, da cognição. Neste âmbito, a nossa própria ênfase recairá sobre algumas expressões mais contemporâneas destas teorias que fazem da ideia de *autonomia* biológica (do Grego antigo *αὐτονομία*, *norma* + *si próprio*; sig. auto-regulação; independência) um alicerce fundamental para o estudo dos biosistemas e da cognição, sendo que, neste contexto, iremos destacar a qualidade *interactiva* da autonomia biológica.

## 1. A Natureza Interactiva dos Biosistemas

O *Interactivismo* (*Interactivism*) é uma abordagem ao estudo de biosistemas, cognição e consequentes faculdades de *representação* que sublinha o carácter *interactivo* dos biosistemas como uma noção *primitiva* na sua ontologia [cf. e.g. Hooker 2009:515]<sup>6</sup>. Ela começa com Robert Rosen [1970 e 1985], Francisco Varela [1979], e Rodney A. Brook [1991]. É a Varela que devemos o entendimento mais recente da *autonomia biológica*, e a Rosen que devemos a sua formalização enquanto sistemas *antecipatórios*. Estes são sistemas que antecipam certas condições do seu meio ambiente, *interagindo com ele de tal forma que a sua subsistência seja preservada* – esta qualidade é de suma importância, caracterizando todos os sistemas que se dizem *autónomos*.

Contudo, há que sublinhar que a interacção entre os biosistemas e aquele ambiente que lhes é fisicamente exterior é apenas um de dois tipos de interacção relevante para a manutenção da autonomia [Bich & Moreno 2016:13]. Não menos relevante é o carácter *autopoietico* (*autopoietic*) [Maturana & Varela 1972] segundo o qual os sistemas, mediante interacções entre as suas partes, se produzem a si mesmos, atingindo *fechamento organizacional* (*organizational closure*)<sup>7</sup> [Moreno & Mossió 2015:xxvi–xxvii] dentro do limite físico-topológico do sistema. Por

---

<sup>6</sup> Dizemos biosistemas num sentido mais próximo de *indivíduos biológicos* do que de *organismos* porque existem sistemas fechados e com organização biológica/auto-regulação que são compostos por múltiplos organismos (e.g. corais, caravelas-portuguesas, biofilmes, etc.). De facto, podem inclusive existir sistemas com auto-regulação que não são unicamente biológicos (e.g. negócios, cidades, regiões geopolíticas e nações [cf. Christensen & Bickhard 2002:3]).

<sup>7</sup> Agradecemos a Davide Vecchi pela indicação de que o conceito de “fechamento organizacional” é bastante problemático; para perceber porquê basta perguntar como é que um sistema pode «adquirir resultados para um fechamento organizacional futuro que contribuam para manter a autonomia» («*achieve future closure outcomes that contribute to maintaining autonomy*») [Hooker 2009:525] e como podemos determinar quais são estes resultados. Esta questão é especialmente problemática nos casos em que a

consequente, a teoria interactivista agrupa princípios externistas e internistas nas suas explicações, respectivamente.

Faça-se um parêntesis para dar conta doutras terminologias utilizadas para referir o fenómeno de autonomia biológica: de acordo com a cibernética, *auto-regulação*; de acordo com a teoria dos sistemas, *homeostasia* [von Bertalanffy 1968:161] (ainda que *homeodinâmica* seja provavelmente um melhor termo)<sup>8</sup>; em termos puramente físicos, *estabilidade não-estacionária fora do equilíbrio termodinâmico*<sup>9</sup>; e, em termos utilizados pelos autoproclamados interactivistas numa fase mais inicial do seu trabalho, *auto-manutenção recursiva* [e.g. Bickhard, Christensen, Hooker, e Collier]. Consideraremos estes termos como sinónimos de modo a evitar sobre-repetição.

Se a autonomia é uma condição *necessária ou suficiente* à vida (ou à manifestação de *agência* que tipicamente associamos a sistemas vivos [Hooker 2009:517]) ou a *sistemas complexos adaptativos* (*complex adaptive systems*) [Gell-Mann 1994:348 e Holand 2006] é uma questão que deixaremos em aberto, com a adenda de que aceitar esta condição como suficiente poderá acarretar por consequência a aceitação de que alguns autómatos, de que são exemplo robots com a capacidade de se alterar a si próprios e de recarregar as suas baterias, são seres vivos<sup>10</sup>. À parte esta questão, é certo que o conceito nos permite traçar uma distinção entre o que é autónomo e outros tipos de sistemas *não-lineares* e *dissipativos* [cf. Prigogine 1980] que envolvem os *processos químicos* específicos responsáveis pela manutenção da coerência metabólica dos sistemas autónomos (*autocatálise*, *autoinibição* e *catálise cruzada*) [Prigogine & Stengers 1984:153], levando-os ao fechamento organizacional. Dito isto, considere-se um exemplo-tipo da literatura, primeiramente utilizado por um dos fundadores da *Epistemologia Evolucionária*, Donald T. Campbell [cf. 1974, onde é cunhado o termo, e Campbell & Bickhard 2003], com vista a esclarecer o que se entende por autonomia: a vela.

Uma vela acesa segue um «princípio de economia» [Wuketits 1990:99] mantendo-se estável fora do equilíbrio termodinâmico à custa da manutenção de dois processos *cíclicos* simultâneos e interdependentes [Hooker 2009:517] (ou “ciclos de restrição de trabalho” (*work-constraint cycles*), na terminologia de Stuart Kauffman [cf. e.g. 2000:4-5, 72, 83, etc.]). A vela consome *energia ordenada* ou combustível (cera) que é vaporizada no pavio à custa da sua combustão, e

---

autonomia de um sistema é largamente dependente de outro sistema, dos quais simbioses e comunidades de organismos são exemplares. Dado que a nossa ênfase neste ensaio recai sobre a dimensão interactiva e não sobre a dimensão organizacional da autonomia, cremos que esta problemática não afectará grandemente as teses por nós avançadas e defendidas, ainda que mais investigação seja necessária futuramente de modo a propor uma definição mais satisfatória.

<sup>8</sup> Sistemas *estáveis não-estacionários* poderão ainda assim ser descritos fisicamente por estruturas matemáticas, essas, sim, *estáticas*. É apenas quando a descrição versa acerca das estruturas, *i.e.* dos modelos matemáticos dos sistemas alvo, que se pode dizer que os sistemas são estáveis no sentido em que não mudam ao longo do tempo, doutro modo a *persistência diz-se unicamente do sistema dinâmico como um todo*. Assim, seguindo António Damásio [2010] e Terrence Deacon [2011] consideramos que *homeodinâmica* é um melhor termo do que *homeostasia* (*etim.* Grega, *στάσις*) para descrever esta estabilidade.

<sup>9</sup> A estabilidade pode ser caracterizada de várias formas, e.g. estabilidade de Lyapunov; e o comportamento diz-se *não-estacionário* porque há flutuações mesmo nos valores das variáveis mais robustas que constituem o sistema, e.g. de acordo com as relações de reciprocidade de Onsager.

<sup>10</sup> Gostaríamos de apontar que esta consequência, ainda que *prima facie* contra-intuitiva, poderá tornar-se mais atractiva após consideração da possibilidade de existência de formas de vida que não têm por base o carbono. Por outro lado, os vírus são sistemas que só registam actividade metabólica e de auto-manutenção quando invadem biosistemas, utilizando os metabolismos destes para se replicarem, podendo também desafiar o estatuto de condição necessária à vida que aqui atribuímos à autonomia. Por alternativa podemos negar que vírus são seres vivos em todos os estados físicos do seu sistema.

esta por sua vez origina correntes de convexão no ar em torno da chama que permitem ao dióxido alcançar a chama. Com estes dois processos origina-se uma *assimetria* termodinâmica na qual o sistema *dissipa energia em maior estado de entropia (mais desordenada)* para o ambiente [cf. Hooker 2009:517, 2011:44, e Moreno & Mossio 2015:xviii]. Notoriamente, caso algum destes dois processos de interacção com o ambiente sejam interrompidos, a vela irá tender para o equilíbrio termodinâmico. Diz-nos portanto Mark C. Bickhard:

«Existe uma relação interna entre um organismo e o seu ambiente em virtude do organismo ser um sistema fora do equilíbrio termodinâmico. Tais sistemas estão necessariamente em interacção com o seu ambiente de modo a manter a sua condição fora de equilíbrio. Isto é uma necessidade física: é assim que funciona a termodinâmica. Se não houver tal interacção, o sistema tende para o equilíbrio e deixa de existir.»<sup>11</sup> [Bickhard 2003:102]

Porém, esta *relação interna e nomologicamente necessária* entre o sistema e o seu ambiente que lhe permite manter a sua estabilidade fora do equilíbrio termodinâmico é bastante limitada no caso da vela. Porquanto a vela não dispõe de dinâmicas que lhe permitam manter a *recursividade da sua auto-manutenção*, viz. que lhe permitam garantir a persistência dos dois processos cíclicos, não se poderá dizer genuinamente autónoma. Assim, a vela não constitui um caso putativo de sistema autónomo *porque a sua estrutura é estacionária ou não-dinâmica*. Noutras palavras, dada a *invariância estrutural ou organizacional* de uma vela, a vela não passa por alterações nem na sua integridade nem relativas à sua interacção com o seu meio, pelo que a sua “autonomia relativa” depende unicamente do acesso imediato a combustíveis (cera e dióxido). Ao contrário, os sistemas autónomos sofrem alterações na sua estrutura, o que lhes permite manterem-se *viáveis* (outro termo para *estáveis*) *numa maior variedade de condições ambientais*. É isto que nos leva a atribuir-lhes comportamentos típicos de agentes [Bickhard 2001:462]. Ademais, é a capacidade dos sistemas manterem a sua auto-manutenção face a variações ambientais que nos permite determiná-los como *robustos* [cf. e.g. Wagner 2005]. Mas, para melhor compreender esta diferença, sugerimos considerar outro caso canónico da literatura: a bactéria.

Em particular consideremos uma bactéria com quimiotaxia (e.g. *Escherichia coli*) positiva (em resposta a químicos que atraem a bactéria, e.g. glucose) e negativa (em resposta a repelentes) [cf. e.g. Maki *et al.* 2000]. Esta bactéria interage com o meio recorrendo ao uso de um flagelo que desempenha actividades motrizes em resposta a concentrações de químicos no seu ambiente [Bickhard 2003:102 & 2009]. Na esteira de Matthew D. Egbert *et al.* [2010], Leonardo Bich & Alvaro Moreno [2016:16] propõem três modelos possíveis para explicar o comportamento quimiotáctico desta bactéria: o primeiro, mais simples, dá conta de uma taxia inteiramente determinada pelo metabolismo, i.e. não respondendo directamente à *deteção* de pistas (sob a forma de *estímulos/sinais* ou *ruído*) do meio ambiente, mas aos metabolitos presentes no plasma intracelular; o segundo apresenta *regulação por receptores sensórios*, com mediação entre estes e a acção motora por parte do metabolismo celular; o terceiro, e mais complexo, difere do segundo na medida em que não há mediação metabólica entre os receptores e a actividade motriz por parte do metabolismo – diz-se, então, que há *desacoplamento estequiométrico (stoichiometric decoupling)*.

---

<sup>11</sup> «There is an internal relation between an organism and its environment in virtue of the organism being a far from thermodynamic equilibrium system. Such systems are necessarily in interaction with their environment in order to maintain their far-from-equilibrium condition. This is a physical necessity: that is the way the thermodynamics works. If there is no such interaction, the system goes to equilibrium and ceases to exist.»

Deste exemplo devemos reter que o *sucesso da bactéria em permanecer autónoma deriva de ela encontrar as condições ambientais que contribuem para a manutenção da sua autonomia* (no caso concreto, diz-se que há sucesso quando a bactéria sobe o gradiente de glucose no ambiente, que lhe é mais favorável). De acordo com o primeiro modelo, a bactéria responde às condições ambientais através daquilo a que C. H. Waddington chamou de «amortecimento absorvente» (*absorptive buffering*) [1968:14] e Rosen de «estabilidade dinâmica» [1970]<sup>12</sup>, sendo que estas condições exteriores não são detectadas (pelo menos) *directamente* pela bactéria. Neste caso, porque o sistema se limita a sofrer e a reagir a causas ambientais que, do ponto de vista informativo, não passam de ruído, podemos dizer que não há detecção das causas da sua alteração por parte do sistema. O sistema limita-se a ajustes dos seus subsistemas emparelhados, mantendo-se no mesmo *atractor* ou, possivelmente, mudando de atractor. Quando o sistema não consegue “absorver” as perturbações, o atractor colapsa e o sistema desintegra-se. O segundo modelo descreve um sistema que já reage *directamente* às condições do meio e não de acordo com os metabolitos que vai produzindo devido à sua permanência no meio, mas ainda é modelado pelo metabolismo. Esta condição já não se verifica no terceiro modelo, em que as operações do *subsistema regulador das interações com o meio ambiente* não são especificadas nem determinadas pela actividade metabólica da bactéria [Bich & Moreno 2016:15].

## 2. A Natureza Evolutiva dos Biosistemas

Para além da sua autonomia, outra qualidade comum a todos os biosistemas é que estes evoluem (parece ser o caso que todos os sistemas evolutivos estudados pela Biologia são sistemas autónomos). Todavia, a qualidade evolutiva dos biosistemas recebeu estudos muito mais aprofundados na Biologia do que a autonomia, que é um muito recente objecto de análise por parte dos biólogos em geral, não integrando sequer o corpo teórico leccionado em cursos de Biologia. Algo semelhante se passa nos estudos acerca da cognição. A sua esmagadora maioria tem por base teorias evolucionárias externistas (e.g. Darwinismo) ao invés de teorias autonómicas com elementos externistas e internistas. O facto é que a teoria biológica encontra-se longe de estar completa (se é que é sequer possível dizer-se duma teoria ou domínio científico que alguma vez estará completo). Historicamente, foi somente no século XIX que o desenvolvimento da Biologia acelerou, vindo-se a constituir um corpo teórico mais coeso apenas no séc. XX, com a chamada “Síntese Moderna” [cf. Huxley 1942]. Porém, já então se falava de uma *extensão* da síntese moderna [cf. Laland *et al.* 2014 e 2015]<sup>13</sup> capaz de englobar fenómenos biológicos até então desconhecidos e que parecem não aceitar uma descrição satisfatória dentro do enquadramento da Síntese Moderna. A autonomia biológica é um desses fenómenos e dedicaremos este capítulo a uma breve investigação da relação entre teorias autonómicas e teorias evolucionárias.

### 2.1. Evolução: A Dimensão Histórica da Autonomia

É comum pensar a evolução de outros sistemas para além dos biológicos. Num extremo, até já foi teorizada a evolução Darwiniana do Universo [e.g. Smolin 1997]. Mais interessante para nós é a consideração de que os sistemas biológicos actuais tomam as suas qualidades apenas

---

<sup>12</sup> Aqui divergimos do uso terminológico de Rosen na medida em que consideramos que todos os modelos descrevem formas de *estabilidade dinâmica*.

<sup>13</sup> Para uma breve introdução ao debate em torno da síntese alargada cf. Laland *et al.* [2014] & [2015].

como resultado de um processo evolutivo pré-biótico. Esta hipótese é comumente chamada de “abiogénese” e pode ser resumida simplesmente como a ideia de que a vida advém da não-vida. Esta hipótese é relevante para nós porque a autonomia é usualmente trabalhada como a propriedade a ser modelada pelas teorias que procuram explicar a *origem* da vida e não somente a sua *transformação*, que é o que sucede com a maioria das teorias evolutivas (e.g. Neodarwinismo), que circumscrevem a evolução biológica à «descendência com modificação» (*descent with modification*) [Darwin 1859]. Em verdade, muitos dos fenómenos que se estudam na teoria autonómica são fenómenos naturais de auto-organização espontânea e não-biológica (e.g. envelhecimento de proteínas) ou de processos que evoluem para estados ordenados (ou atractores) [cf. Brandon 2006:271].

O estudo destes fenómenos é hoje designado por “Química de Sistemas” [e.g. Pascal & Pross 2016]. Mais recentemente, esta área tem-se dedicado ao estudo duma outra qualidade (aparentemente) comum a todos os sistemas vivos: a sua multiplicação por auto-replicação ou reprodução. Alguns autores a trabalhar na área acreditam que a vida se resume a uma forma mais complexa destas mesmas dinâmicas pré- e proto-biológicas<sup>14</sup> [Brandon *op cit.*; cf. Nozick 1993:114-7], enquanto outros acreditam que a “sobrevivência do mais apto” é um caso especial duma lei mais geral da sobrevivência do mais estável<sup>15</sup> [Dawkins 1976:12; e.g. Lotka 1922b:151; cf. Toman & Flegr 2017:30]. Outros ainda crêem que todos os fenómenos tipicamente evolutivos se deixam de algum modo explicar pelas dinâmicas autonómicas dos biossistemas [e.g. Rosslenbroich 2014:225-9 e Moreno & Mossio 2015:xxi], e que a evolução equivale a um processo no sentido de um aumento da autonomia [Rosslenbroich 2014:232]. Para estes últimos autores, a evolução pode então pensar-se como a “dimensão histórica da autonomia” [Moreno & Mossio 2015:§5].

O nosso parco conhecimento em matérias de tamanha complexidade não nos permite ajuizar este *slogan* satisfatoriamente; contudo, há que reconhecer que alguns fenómenos tipicamente estudados pelos teóricos evolucionários se tornam mais *coerentes* quando enquadrados pela perspectiva autonómica. A título de exemplo, a natureza *organizacional* e *interactiva* dos biossistemas reflecte-se em consequências como a *variação genética* e a *organização ecológica* para o biosistema, respectivamente [Hooker 2015:v], deste modo condicionando fortemente os possíveis arranjos *viáveis* dos diferentes biossistemas e, consequentemente, a sua *adaptabilidade* (*adaptiveness*) [Christensen & Hooker 2000:136-7].

Noutro aspecto, a teoria autonómica permite-nos estabelecer hipóteses nulas acerca do processo evolutivo contra as quais podemos testar várias teorias evolutivas. Neste sentido, é possível argumentar, com Alan Turing [1952], Kauffman [1993], ou Robert Brandon [2006:278], entre outros, que sistemas autónomos exibem propriedades dinâmicas ao longo do tempo simplesmente em virtude de se encontrarem fora do equilíbrio termodinâmico. Estas propriedades são assim passíveis de explicações independentes do *processo de selecção natural*, i.e. de descrições evolutivas dadas em termos de *sobrevivência e reprodução diferencial de indivíduos em virtude dos seus fenótipos* (de acordo com uma definição familiar do processo de selecção natural). Por outro lado, fortalecendo a tese de Moreno & Mossio [2015] de que a evolução é a “dimensão histórica da autonomia”, há elementos da teoria autonómica que parecem *integrar* sem

---

<sup>14</sup> Devemos a Maria Rosa Paiva ter-nos chamado a atenção para a possibilidade de haver “selecção” (*sorting*) de estruturas estáveis sem hereditariedade, de acordo com um *princípio de persistência*, do qual a selecção natural é apenas um caso específico – o de sistemas com *estabilidade dinâmica* (por oposição a *estática*) [cf. Toman & Flegr 2017].

<sup>15</sup> «Darwin’s “survival of the fittest” is really a special case of a more general law of survival of the stable (...). The earliest form of natural selection was simply a selection of stable forms and a rejection of unstable ones. There is no mystery about this. It had to happen by definition.» [Dawkins 1976:12 & 13]

grandes dificuldades a teoria evolutiva, e vice-versa, no sentido em que não se contradizem e respeitam os modelos causais umas das outras [cf. Mitchell 2004:84-6 para pormenores acerca do processo de integração epistémica]. No sub-capítulo que se segue iremos considerar um elemento de teorias evolutivas susceptível de integração na teoria autonómica, nomeadamente, a definição de *fitness* na sua relação com a *teoria da selecção natural*.

## 2.2. Autonomia e Fitness

A *Teoria da Selecção Natural* é uma teoria evolutiva avançada em paralelo por Charles Darwin [cf. 1858 e 1859] e Alfred Russel Wallace [cf. Dawkins 2003] com vista a explicar a *origem* e *diversidade* das *espécies*, mas também a *preservação das espécies na luta pela vida* (conforme o sub-título d'*A Origem das Espécies* [Darwin 1859]). Surripiando caminho, em termos muito gerais e abstractos, a teoria propõe, entre outras coisas, uma medida para o sucesso evolucionário pela avaliação da *relação entre os organismos e o seu ambiente*, entendendo que as características que conferem sucesso aos seus portadores seriam aquelas que se encontrariam nos indivíduos vivos e, com maior relevância, teriam maior probabilidade de se encontrar nos indivíduos futuros.

Semelhantemente ao que se passa na teoria interactivista, a *interacção* também é um primitivo ontológico da teoria de selecção natural, mas enquanto a teoria interactivista reúne elementos externistas e internistas nas suas explicações, a interacção em causa na teoria da selecção natural é unicamente exterior, dizendo respeito à interacção causal dos organismos com o seu ambiente, exposta e avaliada em termos da sua *adaptação*.

Diz-nos a teoria da selecção natural que *os organismos adaptados são aqueles que estão aptos a viver num determinado ambiente*, sendo que esta *adaptação* é uma vantagem conducente ao sucesso. Ela é explicada pelas mudanças impostas aos organismos pelo ambiente, mediante um moroso processo causal de selecção natural operando sobre organismos fenotipicamente distintos. A adaptação distingue aquelas *características* que os organismos que se encontram melhor adaptados têm das que se encontram nos pior adaptados a um determinado ambiente<sup>16</sup>. Estas características são transmitidas hereditariamente e a sua presença ou ausência introduz um critério causal que permite discriminar positivamente os organismos que melhor subsistem e se replicam e negativamente aqueles que não têm sucesso em manter-se vivos (pelo menos) até se multiplicarem.

A medida do sucesso evolucionário veio a designar-se comumente por *fitness*<sup>17</sup>, mas enquanto o conceito de fitness surge já nas últimas páginas da primeira edição d'*A Origem* [1859:472] [cf. Krimbas 2004:186 e n.1], terá sido Herbert Spencer a popularizar o seu uso a partir da famosa expressão “sobrevivência do mais apto” (*survival of the fittest*) [1864-7, 1:444-

---

<sup>16</sup> Todavia, hoje em dia há quem argumente que a fitness pode ser uma propriedade não só de características, mas de indivíduos por inteiro, ou mesmo de um grupo [cf. Rosenberg & Bouchard 2015].

<sup>17</sup> Na ausência de uma tradução portuguesa apropriada para este termo continuaremos a utilizar o termo Inglês sem destaque em itálico. Noutro ponto, concordamos com Brandon [1990] para quem *fitness* é equivalente a *adaptação* (*adaptedness*).

5 & 1898:311] [cf. Krimbas 2004:n.2] que o próprio Darwin veio mais tarde a adoptar (por recomendação do seu colega Wallace [1866]) como alternativa ao uso da expressão “selecção natural” para explicar as adaptações das espécies<sup>18</sup>.

Segundo Kent A. Peacock [2011:100], não só Wallace, mas também Spencer tinha já pensado em 1852 os germes da teoria da selecção natural, pois argumentou que os organismos que aumentam em número são aqueles cujo «poder de auto-preservação é maior» [Spencer 1852:500]<sup>19</sup>. Para o nosso presente propósito é notório o quanto Spencer se aproximou da teoria autonómica (ainda que Spencer não tivesse em mente os desenvolvimentos mais recentes em torno do conceito) e cremos que isto se pode dizer de muitos dos primeiros teóricos da evolução por selecção natural, dada a ênfase que colocavam nas relações entre os organismos e o seu ambiente. P.ex., se considerarmos o trabalho de Spencer e de Darwin, em particular, verifica-se que o conceito de fitness que ambos tinham em mente retomava a semântica *vernacular* do inglês *fit* (*encaixe*). A parelha entre uma chave e uma fechadura permite-nos, por intermédio de uma metáfora [de Ariew & Lewontin 2004:348], dar conta deste significado mais intuitivo de fitness. Esta metáfora dá-nos a entender que o que está em causa são as relações causais entre as propriedades dos organismos e a estrutura do seu ambiente (da chave e da fechadura), razão pela qual é costume referir esta noção de fitness como *ecológica* (outros autores chamam-lhe ainda “*fitness Darwiniana*”). Esta noção de fitness ecológica enfatiza a capacidade dos organismos responderem aos desafios ou «*problemas de design*» (*design problems*) [cf. Dennett 1995] que o ambiente coloca à sua sobrevivência, e que é bem realçada na seguinte formulação de Alexander Rosenberg & Frédéric Bouchard:

«x é mais fit que y se, e somente se, os traços de x lhe permitirem resolver os “problemas de design” colocados pelo ambiente melhor que os traços de y.» [2015:3]<sup>20</sup>

Um problema imediato desta definição, que será relevante para um momento mais tardio do nosso ensaio, é que não é imediatamente claro quais são as características do ambiente que são relevantes para a sobrevivência de determinado indivíduo – chamamos a isto “*o problema do ambiente de referência*”, assim cunhado por Marshall Abrams [2009] [cf. ainda Rosenberg & Bouchard 2015:3-4, Lewontin 1978, e Bouchard 2008:561]. É ainda de realçar que, de acordo com esta leitura, fitness é uma qualidade que só pode ser avaliada *qualitativamente relativa a um ambiente específico*. Robert D. Rupert diz por isso que é *eco-dependente* [Rupert 2009:79-80]. Preferimos todavia a terminologia de Brandon, o qual nos diz que a fitness é uma propriedade *extrínseca* [Brandon 1990:39] (por oposição a uma propriedade *intrínseca*, definida como uma propriedade que existe ou se manifesta *independentemente* daquilo que se passa para além do seu portador [McKittrick 2003]).

Notoriamente, estas características reflectem-se na teoria autonómica. Podemos perguntar quais são as condições ambientais nas quais as interacções dos biosistemas são conducentes à preservação da sua autonomia, e a avaliação da autonomia de um biosistema é relativa a um determinado ambiente com propriedades devidamente especificadas. Na possibilidade de avaliar a fitness não só de indivíduos, mas de características de indivíduos em concreto, repare-se que, no contexto da teoria autonómica, uma determinada característica nunca pode receber

---

<sup>18</sup> Wallace recomendou o uso da expressão «survival of the fittest» a Darwin com a justificação de que o emprego da expressão alternativa cunhada por Darwin “selecção natural” confundiu muitos leitores a interpretar a Natureza como um agente intencional que determina o *design* das espécies [cf. Wallace 1866].

<sup>19</sup> «the power of self-preservation is the greatest».

<sup>20</sup> «x is fitter than y if and only if x's traits enable it to solve the “design problems” set by the environment more fully than y's traits do.»

uma avaliação completa relativamente à sua contribuição para a manutenção da autonomia sem por isso pesá-la no complexo de interacções em que se encontra com outras características interiores ao próprio sistema – a dimensão organizacional ou internista da teoria autonómica. Estes dois elementos combinados, o ecológico e o interior ou organizacional, dão lugar àquilo que poderemos designar por “holismo da fitness/autonomia”.

Foi somente com os desenvolvimentos da Genética e da Biologia Populacional (em especial) durante as décadas de 20-40 que outros elementos, entre os quais o reprodutivo, tomaram primazia nas definições de fitness [cf. Sober 2001:309]<sup>21</sup>, ainda que o elemento reprodutivo esteja já presente desde a primeira publicação da teoria da evolução por selecção natural. O diferendo reside em que nestas primeiras teorias o elemento causal ou ecológico dos fenótipos precedia o reprodutivo nas explicações de sucesso evolucionário, enquanto em algumas formulações do séc. XX o elemento fenotípico chega a desaparecer em detrimento duma interpretação meramente estatística das dinâmicas populacionais [cf. Ariew & Lewontin 2005]<sup>22</sup> de acordo com a qual um aumento na representatividade de um gene equivale a um aumento na fitness, independentemente dos detalhes acerca da sua contribuição causal ou ecológica (para o seu fenótipo). É esta interpretação que exclui os fenótipos que pretendemos aqui criticar, enfatizando o carácter causal e ecológico *ou* interactivo dos biosistemas, *i.e.* a fitness Darwiniana, uma vez que, na sua presente forma de apresentação, as teorias autonómicas tomam uma descrição sobremodo causal e que os próprios teóricos da autonomia criticaram a ênfase da teoria genético-populacional na reprodução, negligenciando o carácter interactivo dos seres vivos [cf. e.g. Varela 1979:5]<sup>23</sup>.

A nossa crítica é simples e tem por base o próprio conceito de autonomia conforme previamente definido: a autonomia não exige a existência de reprodução (ou mesmo de multiplicação) dos biosistemas ditos autónomos. Então, *a fortiori*, a compatibilização das teorias requererá a formulação de uma medida de sucesso evolucionário que não dependa somente da multiplica-

---

<sup>21</sup> O conceito de *fitness* tem sido largamente discutido, em particular nas últimas duas décadas de anos [cf. Rosenberg & Bouchard 2015]. Autores como “o último” Karl Popper e Brandon [cf. 1990:139-40] pensaram inclusive que as complicações derivadas das tentativas de definir o conceito de fitness requerem que repensemos o estatuto epistémico da teoria da selecção natural [cf. Rosenberg & Bouchard 2015:15] (Popper, a título de exemplo, acabou por pensar no Darwinismo, *i.e.* na teoria da selecção natural que se segue das interpretações de Darwin, como um *programa de investigação metafísico* e não como uma *teoria científica* propriamente dita [cf. 1976:168], vindo mais tarde a moderar a sua perspectiva [1978:346]). Outros autores, como Ariew e Lewontin [2004] concluíram que não há nenhum conceito de fitness que unifique todos os diferentes significados de fitness, optando no invés por uma abordagem pluralista.

Aqui deixaremos de lado alguns dos problemas usualmente dirigidos às definições de fitness (e.g. falsificação, aparente tautologia [cf. Rosenberg & Bouchard 2015]). Iremos inclusive deixar de lado alguns dos problemas relacionados com a própria definição de fitness que iremos utilizar [*idem*]. Iremos tratar a propriedade fitness como uma *propensão*, *i.e.* uma *disposição probabilística*, com base nas propriedades causais dos traços. Assim, a fitness pode não se actualizar (e.g. o traço pode não se multiplicar na população apesar de ser potencialmente mais vantajoso). Note-se que esta interpretação é várias vezes invocada na tentativa de dissolver o aparente carácter tautológico do conceito de fitness [Sober 2001].

<sup>22</sup> Dado que a teoria termodinâmica é sobremodo estatística e que o mesmo se passa com as teorias genéticas populacionais, talvez a possibilidade de uma harmonização destas teorias não esteja epistemicamente distante face ao actual estado da arte.

<sup>23</sup> Especialmente claro no seguinte excerto: «evolutionary thought, through its emphasis on diversity, reproduction, and the species in order to explain the dynamics of change, has obscured the necessity of looking at the autonomous nature of living units for the understanding of biological phenomenology. Also I think that the maintenance of identity and the invariance of defining relations in the living unities are at the base of all possible ontogenetic and evolutionary transformation in biological systems.»



ção de indivíduos. Isto torna-se mais claro se considerarmos que os sistemas autónomos apresentam *adaptabilidade* às condições do meio, *i.e.* sofrem de alterações nas suas dinâmicas provocadas por causas ambientais de tal modo que se mantêm viáveis, e que estas dinâmicas evolutivas não necessitam de envolver reprodução. Um exemplo *hipotético* dá conta disto: pense-se um único indivíduo que *interage com, e responde adaptativamente a, um meio* sem nunca se multiplicar, subsistindo nesta condição ao longo de várias eras geológicas. Mesmo não se reproduzindo, é indiscutível que este indivíduo hipotético está não obstante a “fazer alguma coisa bem” – para usar uma expressão de Bouchard [2008:569] – e, portanto, face à necessidade de explicar o sucesso evolutivo deste indivíduo, há que adoptar uma definição de fitness que não dependa da reprodutibilidade.

A relevância deste exemplo é limitada dado que não se sabe se existe ou não tal indivíduo e uma vez que todos os *indivíduos biológicos* conhecidos parecem multiplicar-se. Posto isto, não é de surpreender que autores como Peter Godfrey-Smith acreditem que ainda que seja possível proceder a uma descrição de mudanças em colecções de coisas que não se reproduzem utilizando para esse efeito o discurso da selecção natural o resultado seria muito artificial e pouco interessante [2009:40]. Samir Okasha [2006:214] chama “*fraca*” a esta forma de selecção natural. Podemos dizer que estes autores são cépticos relativamente à virtude epistémica de teorias selectivas acerca de sistemas que não se reproduzem. Outros consideram mesmo que «[a] Biologia tem início quando a reprodução começa»<sup>24</sup> [Brandon 2008] e que este exemplo hipotético não representa, como tal, um caso biológico.

Todavia, nós acreditamos que os aspectos ecológicos da fitness continuam relevantes dado que existem indivíduos biológicos que desafiam as subseqüentes definições de fitness. Para dar conta destes aspectos num contexto compatível tanto com a teoria autonómica como com a teoria evolucionária, ignoremos por momentos a *possibilidade* do mecanismo de selecção natural actuar em populações de indivíduos que não se reproduzem nem estão em competição uns com os outros (nomeadamente, por anagénesse), como demonstrado por Ford Doolittle [2014:418, n.2, e p.420], e consideremos unicamente indivíduos que se multiplicam. Para este efeito, e antes de prosseguir para considerações acerca de indivíduos biológicos específicos, consideremos um tipo especial de sistemas autónomos descritos por Rosen e a que este chamou de “*sistemas-(M, R)*” (de *metabolic repair*) [para uma versão detalhada cf. Rosen 1985:203-7].

Sistemas-(M, R) são tais que as interacções de alguns dos seus conjuntos autocatalíticos produzem inadvertidamente outros sistemas, *i.e.* são auto-replicantes. Rosen propôs estes sistemas com o intuito de modelar a possibilidade de sistemas autorregeneradores se replicarem no habitual processo de autorregeneração que leva à manutenção da recursividade dos sistemas autónomos. Nas palavras de Wayne D. Christensen, «se um (...) sistema auto-organizador satisfizer certas condições, nomeadamente, a de que a correspondência biunívoca entre o *output* e o *input* metabólico utilizável é invertível, então o sistema-(M, R) irá produzir como *output* outro sistema-(M, R)»<sup>25</sup> [1996:308]. Como resultado, estes sistemas «exibem dois tipos de estabilidade temporal. Primeiramente, a sua organização é tal que são auto-reparadores, e como tal são estáveis *durante o desenvolvimento*; em segundo lugar, um subconjunto destes sistemas é capaz de se auto-replicar, e como tal de demonstrar estabilidade *inter-geracional*» [*ibidem*]. Neste contexto, estabilidade inter-geracional fora do equilíbrio termodinâmico torna-se sinónimo de

---

<sup>24</sup> «Biology starts when reproduction begins.»

<sup>25</sup> «if a (...) self-organising system satisfies certain conditions, namely that the mapping from output to usable metabolic input is invertible, then the (M, R)-system will produce as output another (M, R)-system (...) / Thus (M, R)-systems exhibit two kinds of temporal stability. Firstly their organization is such as to be self-repairing, and hence is developmentally stable. Secondly, a subset of them is capable of self-replication, and hence shows inter-generational stability.»

evolução filogenética por replicação (ainda que não necessariamente por reprodução diferencial, dado que estes sistemas produzem clones de si próprios, pelo que o tipo de replicação aqui em causa seria mais próximo daquilo a que, nos organismos, chamamos de reprodução assexuada).

Este modelo de sistemas-(M, R) permite-nos trabalhar a teoria autonómica em paralelo com a teoria evolutiva, dado que estes sistemas se replicam necessariamente no decurso dos seus processos auto-regenerativos. Com estes sistemas, podemos então dispensar a possibilidade do anterior exemplo hipotético e inclusive aceitar, com Brandon [*Op. Cit.*], que a replicação é *conditio sine qua non* do sucesso evolutivo, ou pelo menos que «[a] Biologia neste planeta começou quando um certo tipo de auto-manutenção se tornou possível»<sup>26</sup> [Peacock 2011:101]. Ademais, como a auto-replicação permanece uma forma eficiente de dissipar energia [cf. England 2013], podemos fazer sentido desta estratégia à luz da termodinâmica dos sistemas autónomos. Dito isto, o que gostaríamos de demonstrar aqui é que existem biosistemas para os quais a reprodutibilidade se afigura como um critério insatisfatório da medida da sua fitness e que, num extremo, a reprodução é apenas uma estratégia possível para a preservação da subsistência dos biosistemas<sup>27</sup>, podendo não ser muito relevante [cf. e.g. Bouchard 2008].

Seguimos na senda dos críticos das definições afins ao paradigma populacional ou Malthusiano que, conforme mencionado acima, culminou com os estudos genéticos e biomatemáticos da primeira metade do séc. XX. Nas palavras de Elliott Sober [1980], estes estudos matemáticos (e.g. Theodosius Dobzhansky) foram acompanhados por um desinteresse pelas qualidades dos indivíduos que, na interpretação ecológica e causal de fitness, explicam o seu sucesso evolucionário. Porém, são inúmeros os indivíduos biológicos cuja fitness não se deixa entender sem estes aspectos, de que são exemplo violetas, árvores (e.g. o álamo-trémulo, *Populus tremuloides*), corais, caravelas-portuguesas, e fungos (e.g. *Armillaria ostoyae*) [Ariew & Lewontin 2004 e Bouchard 2008 e 2011]<sup>28</sup>. Algo comum a todos os indivíduos biológicos destas espécies é que são na realidade constituídos por múltiplos organismos mais ou menos *estrutural e funcionalmente integrados*, de tal forma que a sua fitness só pode ser atribuída ao sistema como um todo, o qual apresenta propriedades fenotípicas emergentes, e não aos organismos individualmente. Estes, nas palavras de Peacock, não constituem indivíduos biológicos *discretos* [Peacock 2011:101]. Mas o que é mais interessante nestes indivíduos é que raramente se reproduzem com sucesso, tendendo antes a produzir clones de si próprios que *continuam integrados* no organismo que lhes dá origem e, por consequência, no indivíduo como um todo, *viz.* o indivíduo aumenta em expansão/massa sem aumentar em número.

Considere-se o caso do álamo-trémulo em particular. Estima-se que um determinado espécime de álamo-trémulo na América do Norte tenha várias dezenas de milhares de anos e ocupe

---

<sup>26</sup> «Biology on this planet started when a certain type of self-maintenance becomes possible in dissipative systems constructed out of carbon-chain molecules.»

<sup>27</sup> É ainda de notar que a aceitação desta condição acarreta, por implicação, que uma determinada característica que contribui para a viabilidade do sistema e que, simultaneamente, o torna estéril, tem um valor de fitness igual a zero [Sober 2001:309].

<sup>28</sup> André Ariew & Richard Lewontin [2004:360] crêem que estes problemas que se seguem da aplicação dos conceitos tradicionais de fitness a indivíduos biológicos atípicos deve-se a uma leitura Malthusiana da teoria evolucionária segundo a qual é relevante o número populacional para a contabilidade dos recursos disponíveis em determinado nicho ecológico quando este é um *proxy* para os recursos utilizados. É, pois, curioso, que o conceito genético de fitness tenha sido desenvolvido em simultâneo com o estudo de indivíduos biológicos que, devido ao seu igual tamanho, poderão constituir um *proxy* para os recursos utilizados (e.g. *Drosophila melanogaster*). Também relativamente a este aspecto o caso humano é desafiante dada a discrepância registada na exploração de recursos (e.g.) entre indivíduos provenientes de diferentes estados ou classes sociais.

uma área de várias centenas de hectares [cf. Bouchard 2008 & 2011]. Trata-se, pois, de um indivíduo com bastante sucesso evolutivo no que toca ao critério da viabilidade. Mas o que é mais interessante acerca deste indivíduo diz respeito às suas estratégias adaptativas que lhe permitiram manter-se viável durante tanto tempo.

Este indivíduo (o arvoredor) é constituído por vários clones integrados. Um clone ou uma semente (produtos da reprodução sexual) que se separa do indivíduo integrado não terá acesso aos nutrientes partilhados pelos clones integrados através da rede de raízes (ainda que o seu ambiente selectivo seja relativamente semelhante) [Bouchard 2008:564-5]. Podem também dar-se casos em que, porventura por causa do vento, correntes hídricas, ou outros factores de dispersão, as sementes não partilham da *herança ecológica* dos seus progenitores, sendo que o novo ambiente que encontram pode não ser favorável ao seu sucesso, e a *incerteza* relativamente às condições ambientais pode não compensar a estratégia reprodutiva [Van Valen 1989:5]. Além do mais, o arvoredor integrado consegue adaptar-se mediante alterações nos seus fenótipos (e.g. mudança da área geográfica ocupada) sem que seja acompanhado por mudanças nas frequências dos alelos [Bouchard 2011:108-10]. E talvez com maior relevância para a presente crítica, *os clones integrados competem com a progénie sexual* e tendem a ganhar porque são melhor alimentados através da rede de raízes, o que resulta na eliminação da sua descendência sexual [Bouchard 2011:112]. A todas estas estratégias problemáticas para a definição NeoDarwinista de fitness, Ariew & Lewontin acrescentam ainda uma dúvida: um tipo de indivíduo que ocupa um vasto território tem mais ou menos sucesso que um tipo que reproduz vários milhares de indivíduos que ocupam uma menor expansão de território e que vivem durante menos tempo [Ariew & Lewontin 2004:360]?

Peacock declara ainda outras preocupações com relevância para o caso do *H. sapiens*. Pergunta-nos Peacock, aqui parafraseado: quão fit é um organismo que se reproduz mas garante que as condições ambientais futuras serão tais que a sua progénie *não* poderá dar continuidade à linhagem filogenética? A título de exemplo considere-se um cenário de *risco existencial* [Bostrom 2002] no qual seres humanos esgotam recursos essenciais e cobrem a Terra de cinza nuclear após uma longa Guerra Mundial ou então um cenário de alterações climáticas induzidas pelo *H. sapiens* e com consequências igualmente desafiantes. Se o único tipo de adaptação fosse relativo a condições ambientais *locais e imediatas*, cegas a pressões evolutivas de longo prazo, seria fácil imaginar a vida a auto-obliterar-se em poucas gerações sob dinâmicas típicas de uma «tragédia dos comuns» Hardiniana a uma escala planetária<sup>29</sup> [Peacock 2011:103, e p.101; cf. Hardin 1968, e Corning & Szathmáry 2015]. Indivíduos com tais comportamentos não terão grande sucesso evolucionário.

Curiosamente, o *H. sapiens* partilha algumas semelhanças com o álamo-trémulo. Apesar de existirem poucos indivíduos humanos a partilharem por inteiro o genoma uns dos outros (nomeadamente, gémeos monozigóticos ou univitelinos), também o humano apresenta comunidades *integradas* nas quais se dá a partilha de recursos e, com o actual estado de conectividade dos mercados, esta integração não é apenas local (como no caso do álamo-trémulo), mas global (e.g. podemos transportar recursos para os nossos antípodas). Assim, também não é possível comparar a fitness da maioria dos humanos independentemente desta *conectividade* (que varia

---

<sup>29</sup> «Recall Gould's explication of fitness as Darwin apparently understood it, as adaptation to the "immediate, local environment". If this were the only kind of adaptation of which life was capable, it would have long ago obliterated itself in a planetary-scale Hardinian "tragedy of the commons"».

de *tamanho*<sup>30</sup>) sob pena de ignorar diferentes estratégias evolucionárias nitidamente relevantes para o sucesso dos humanos [cf. Bouchard 2008:565, e Nowak, Tarnita & Wilson 2010]<sup>31</sup>.

É atendendo a estas e outras complicações que Bouchard conclui que o número de descendentes não é necessário para avaliar a fitness de um indivíduo: o destino dos vários componentes ou partes do indivíduo (no caso do álamo-trémulo, os clones integrados) parece bastar para medir a adaptação destes sistemas [Bouchard 2008:565]. Faça-se a advertência: a lição a retirar destes exemplos não é que a reprodução sexuada seja irrelevante ou que o modelo ortodoxo não consiga explicar a fitness de reprodutores assexuados, mas que organismos estruturalmente coesos podem não se reproduzir e ainda assim ter muito sucesso evolucionário [Bouchard 2008:564]. Noutras palavras, como estes indivíduos parecem estar a “fazer alguma coisa bem”, *i.e.* têm sucesso evolucionário fruto da selecção natural, para avaliar a sua fitness não podemos utilizar uma medida estatística que se cinge à contagem de descendentes deixados pelos indivíduos.

Face a estes inúmeros problemas (entre outros [cf. Nanay 2011 & Bourrat 2014]) que desafiam a definição de fitness da Genética Populacional, foram propostas alternativas com vista a avaliar o sucesso de todos os sistemas evolutivos com respeito pelas suas diferentes estratégias evolutivas. Não nos cabendo aqui uma última palavra relativamente a este problema é ainda assim possível verificar que algumas definições se encontram em maior sintonia com a teoria autonómica – e talvez se possa inclusive falar de consiliência (*consilience*) [cf. Wilson 1998].

Comecemos por Ariew & Lewontin [2004:360], que consideraram o sucesso evolucionário diferencial em relação próxima com o controlo de recursos disponíveis, *i.e.* quanto acesso os organismos têm a *energia livre* (*free energy*) ou *energia útil* (*useful energy*) do ambiente, verificando que em alguns casos o número populacional pode ser um *proxy* fiável para esta medida<sup>32</sup>. Bouchard [2008] dá conta que esta definição é próxima da de Leigh Van Valen, que, apesar de ter diferentes preocupações em mente (sobretudo ecológicas), chega a uma definição de fitness enquanto «expansão diferencial esperada»<sup>33</sup> [Van Valen 1989:7], entendendo «a fitness realizada ou momentânea de um indivíduo ou de uma população como a quantia de recursos regulatórios, em especial energia, que controla»<sup>34</sup> [1975:267]. Bouchard observa ainda que estas definições traduzem-se por vezes em maiores números reprodutivos ao invés de expansão dum só indivíduo (como no caso do álamo-trémulo citado acima) [cf. Bouchard 2011:109].

Mas Peacock verifica que o aumento populacional ou em expansão não é necessário para que se possa dizer de uma população ou indivíduo que controle maior quantidade de energia livre (mantendo-se em mesmo número, com a mesma frequência de genes, ou tamanho) que é

---

<sup>30</sup> Adoptamos aqui a terminologia da Teoria de Grafos.

<sup>31</sup> Autores como Edward O. Wilson dão um passo mais longe e questionam a própria ontologia da espécie humana relativamente aos elevados graus de interdependência que caracterizam as suas sociedades, observando que há um limite de integração para além do qual deixa de fazer sentido considerar os organismos individualmente [cf. Wilson 1974:54, 1971:469, e Nowak, Tarnita & Wilson 2010]. Deixaremos de lado estas questões pouco ortodoxas e alvo de muita crítica pós-2010.

<sup>32</sup> Defendendo uma abordagem pluralista, Ariew & Lewontin fazem ainda outras sugestões para revisões às definições de fitness. P.ex.: «As an alternative it would be possible to build a dynamical theory that used not numbers of denumerable objects, but a continuous measure of occupancy of the external world, for example the total amount of living protoplasm belonging to a given type.» [Ariew & Lewontin 2004:360]

<sup>33</sup> «Rather than saying that natural selection is expected differential reproduction, we should say that it is expected differential expansion.»

<sup>34</sup> «The momentary or realized fitness of an individual or population as the amount of its ultimately regulatory resources, especially energy, that it controls.»

mais fit que outra que não exerce controlo sobre a energia livre dos ecossistemas (não sofrendo também outras alterações em número ou expansão). E isto porque a primeira população ou indivíduo protege ou aumenta activamente a sua probabilidade de sobrevivência a longo prazo e a(o) segunda(o) não (o que no caso do *H. sapiens* poderá traduzir-se no cultivo de autotrófos de que se alimenta e dos quais depende para respirar por oposição ao simples consumo de autotrófos sem a sua reposição por cultivo) [Peacock 2011:102]<sup>35</sup>. Estes exemplos encontram-se em clara sintonia com a teoria autonómica, de acordo com a qual os biosistemas que permanecem autónomos agem «no seu ambiente para promover a sua própria manutenção»<sup>36</sup> [Moreno & Mossió 2015:90–1] (e Peacock [2011:101] diz-nos isto mesmo, ainda que não pareça estar familiarizado com a literatura em torno da teoria autonómica<sup>37</sup>).

Contudo, Rosenberg & Bouchard [2015:18-9], Bouchard [2011] e Ford Doolittle [2014] concluem apontando todas estas estratégias dominantes como subservientes duma outra definição de fitness, começando por demonstrar que constituem estratégias viáveis, mas não necessárias, havendo ainda outras alternativas. Para englobar todas estas possibilidades concluem que o que é maximizado aquando da maximização da fitness é a *persistência diferencial das linhagens* (filogenéticas) *ao longo do tempo*:

«(a linhagem) X é mais fit que (a linhagem) Y se X tiver uma maior propensão a persistir por uma quantidade de tempo Z que Y.» [Bouchard 2011:113, ou Rosenberg & Bouchard 2015:19]

Onde as linhagens são «entidades biológicas que se alteram indefinidamente ao longo do tempo, como resultado da persistência diferencial dos seus componentes»<sup>38</sup> [Bouchard 2011:113]. Com esta definição, conseguimos substanciar a contribuição essencial das interacções ecológicas para a persistência e em simultâneo dar conta dos casos que desafiam a definição genético-populacional de fitness, *i.e.* independentemente do sucesso reprodutivo [*ibidem*], que, de acordo com os demais exemplos, parece ser uma estratégia contingente e não necessária para a sobrevivência de uma linhagem [Bouchard 2011:114]<sup>39</sup>.

---

<sup>35</sup> Os *H. sapiens* são, ao fim e ao cabo, exemplares construtores de nichos [cf. Odling-Smee *et al.* 2003].

<sup>36</sup> «acts on its environment to promote its own maintenance»

<sup>37</sup> Ainda relativamente a estas abordagens, Peacock [2011] diz-nos que se aproximam ainda doutra anterior definição proposta pelo biólogo teórico Alfred J. Lotka [1922] (a que Peacock se refere como *fitness constructiva* ou *Lotkania* («*constructive or Lotkan fitness*»)) [Peacock 2011:103]), que dá conta da importância de auto-preservação da vida a longo prazo, apontando como mais fit aqueles que contribuem para o aumento do fluxo total de energia livre dos ecossistemas. Atente-se especialmente ao seguinte excerto de Lotka: «But the species processing superior energy-capturing and directing devices may accomplish something more than merely to divert to its own advantage energy for which others are competing with it. If sources are presented, capable of supplying available energy in excess of that actually being tapped by the entire system of living organisms, then an opportunity is furnished for suitably constituted organisms to enlarge the total energy flux through the system. Whenever such organisms arise, natural selection will operate to preserve and increase them. The result, in this case, is not a mere diversion of the energy flux through the system of organic nature along a new path, but an increase of the total flux through that system.» [Lotka 1922a:147]

<sup>38</sup> «[L]ineages as *biological entities that change indefinitely through time as a result of differential persistence of their components*».

<sup>39</sup> Bouchard observa que a avaliação da fitness em termos de persistência diferencial de linhagens ao longo do tempo não pressupõe que as linhagens se reproduzem ao modo dos organismos: uma linhagem pode divergir sem que tal implique que a linhagem materna aumenta a sua fitness enquanto persistência ao longo do tempo. Ademais recordamos que Ford Doolittle demonstrou a possibilidade de persistência não-diferencial de linhagens sob o efeito de selecção natural, ainda que não tenha sido demonstrada a

Todas estas consequências aproximam-nos, *prima facie*, duma definição compatível com as várias tentativas de unificar as teorias termodinâmicas, autonómicas e evolutivas [cf. Bouchard 2011:107 e Toman & Flegr 2017], de tal forma que podemos concluir, com Hooker, que «um organismo está adaptado quando possui um conjunto de traços que satisfazem a sua autonomia no seu ambiente de vida»<sup>40</sup> [2009:527], e que um aumento da estabilidade dos biosistemas relativamente a influências ambientais corresponde a um aumento da sua fitness ecológica [Ros-slenbroich 2014:230] e, conseqüentemente, à persistência da sua linhagem ao longo do tempo. Dito isto, comprometemo-nos de agora em diante com este conceito de fitness.<sup>41</sup>

### 3. A Teoria Interactivista e Evolucionária das Funções Biológicas

O corpo teórico da Ciência permitiu ao *H. sapiens* alterar as condições do mundo que habita e inclusive alterar-se a si próprio, tudo em virtude das suas capacidades cognitivas. Para o seu próprio benefício ou malefício, nem os maiores pessimistas acerca da Ciência poderão negar que ela se repercute drasticamente no nosso modo de vida. Contra a atitude pessimista, parecem-nos que existe alguma relação positiva entre a capacidade de produzir Ciência e o sucesso biológico dos *H. sapiens*, o que é o mesmo que dizer, em expressões de senso-comum, que a Ciência parece desempenhar alguma função biológica (voltaremos a este assunto mais adiante, cf. §V). Mas para podermos concluir que sim ou não é necessária uma teoria de funções biológicas.

É possível agrupar as várias teorias naturalistas de funções em duas grandes famílias: as que partem de intuições relacionadas com o aparente design dos objectos construindo a partir daí teorias gerais de funções (e.g. teorias causais de funções [cf. Couch 2011]) e outras que se esforçam por seguir a própria teoria biológica, nela fundando a teoria de funções (e.g. teorias teleosemânticas etiológicas e interactivistas). De algum modo, estas últimas permanecem todavia ligadas a noções de design, que provavelmente são inalienáveis das nossas teorias de funções, como sugere Philip Kitcher [1993]. Aqui iremos considerar somente estas últimas noções de função que se procuram fundar na teoria biológica. Relativamente a estas, Collin Allen [2003] sugeriu que se pensasse do seguinte modo a relação entre design e função:

«O traço T está naturalmente desenhado para fazer X significa que

---

existência de casos que aceitem esta descrição. Noutro sentido, Rui Diogo relembra-nos que existem muitas linhagens com bastante «"success" (...) in terms of taxonomic diversity, that are far from being optimal eco-morphological matches» [Diogo 2017:2].

<sup>40</sup> «An organism is adapted when it possesses an autonomy-satisfying set of traits in its life-environment.»

<sup>41</sup> A integração da teoria evolutiva e da teoria da autonomia biológica necessitará ainda de um enquadramento no contexto: (1) da teoria de construção de nichos, que diz respeito à capacidade dos organismos alterarem o seu próprio ambiente e, por consequência, a acção da própria selecção natural sobre si mesmos [cf. Darwin 1881b, Levins & Lewontin 1983, Odling-Smee et al. 2003, Day et al. 2003, e Godfrey-Smith 2001a & 2014]; (2) do debate acerca das unidades que são objecto da selecção (e.g. genomas, organismos, grupos, e espécies) e do nível no qual agem as propriedades pelos quais estes objectos são seleccionados [Engels 1876, Darwin 1871, Vrba & Gould 1986, Okasha 2006, Nowak 2006, Jablonski 2008, Sterelny 2012, Corning & Szathmáry 2015, e Lloyd 2017]; e, por fim, (3) da teoria das hierarquias biológicas, segundo a qual existem duas grandes hierarquias biológicas, caracterizadas por critérios mereológicos, uma genealógica (seguindo critérios de replicação de informação genética) e outra económico-ecológica (seguindo critérios de transferência de matéria e energia entre sistemas) [cf. Eldredge 2008, e Tëmkin & Eldredge 2015], pois parece ser o caso que sistemas autónomos formam hierarquias [Christensen & Bickhard 2002], sendo que a preservação de um sistema autónomo pode ser (in)dependente da preservação dum sistema autónomo de ordem superior que este integre ou vice-versa.

1. X é uma função biológica de T e
2. T é o resultado de um processo de mudança de estruturas (anatômicas ou comportamentais) devido à selecção natural que resultou em T estar mais otimizado (ou melhor adaptado) para X do que versões ancestrais de T.»<sup>42</sup>

Noutras palavras, Allen diz-nos que, de acordo com estas teorias, a existência e permanência duma determinada função se deixa explicar causalmente pelo processo de selecção natural. Aqui iremos tentar demonstrar que a teoria etiológica e a teoria interactivista das funções biológicas concordam neste aspecto, mas aplicam diferentemente a teoria de selecção natural nos seus *explananda*. Em particular, iremos procurar demonstrar que, enquanto a teoria etiológica de funções equaciona características funcionais com *adaptações*, a teoria interactivista ou autonómica equaciona características funcionais com *aptações*, sendo que uma «*aptação* (...) [é] um termo inclusivo para qualquer carácter actualmente sujeito a selecção positiva, independentemente de como evoluiu (...), em contraste com o termo mais restrito “*adaptação*” para um carácter que evoluiu por selecção natural para a sua presente função»<sup>43</sup> [Vrba & Gould 1986:219, nossa ênfase]. Ou ainda, na terminologia de Allen, a teoria etiológica é uma «abordagem que “olha para trás” e que analisa a função de um traço apenas em termos daqueles efeitos do traço que, no passado, contribuíram para a selecção de organismos com aquele traço»<sup>44</sup> [*idem*] e a teoria interactivista é uma «abordagem que “olha para a frente” ou disposicional e que analisa a função em termos daqueles efeitos que [a disposição] está disposta a produzir e que tendem a contribuir para a manutenção presente e futura daquele traço numa população de organismos»<sup>45</sup> [*idem*].

Começaremos por expor uma versão breve e simplificada da teoria etiológica, seguindo-a de críticas que se dissolvem com a teoria interactivista com cuja exposição terminaremos este capítulo.

A teoria etiológica terá sido primeiramente proposta por Ruth G. Millikan [1984] e recebeu uma série de retoques ao longo dos anos por parte de Millikan [e.g. 1993] e doutros proponentes [e.g. Papineau 1987, 1993 e 2001, e Neander 1995]. Apesar das diferenças entre os autores, há um aspecto essencial que permanece inalterado e pode ser apresentado com brevidade: uma função de um determinado traço é o seu “propósito evolucionário”, i.e. o efeito produzido pelo traço e que o terá tornado alvo de selecção natural positiva. Daqui conclui-se que a definição etiológica de função biológica é *histórica e não causal*. Millikan explicitou-o: «a “função própria” de uma coisa (...) tem a ver não com os seus poderes [causais], mas com a sua história.»<sup>46</sup> [Millikan 1984:17]<sup>47</sup>.

A caracterização histórica de funções da teoria etiológica tornou-a alvo de inúmeras críticas. Bickhard, p.ex., observou que a história só tem consequências causais porquanto causas distais

<sup>42</sup> «*Trait T is naturally designed to do X means that: / 1. X is a biological function of T and / T is the result of a process of change of (anatomical or behavioral) structure due to natural selection that has resulted in T being more optimal (or better adapted) for X than ancestral versions of T.*»

<sup>43</sup> «*aptation (...) an inclusive term for any character currently subject to positive selection regardless of how it evolved (...) in contrast to the narrower term adaptation for a character evolved by natural selection for its current function*»

<sup>44</sup> «*backward-looking approach that analyses the function of a trait only in terms of those effects of the trait which have in the past contributed to the selection of organisms with that trait.*»

<sup>45</sup> «*a dispositional or forward-looking approach that analyses function in terms of those effects it is disposed to produce that tend to contribute to the present or future maintenance of the trait in a population of organisms.*»

<sup>46</sup> «[T]he “proper function” of a thing (...) has to do not with its powers but with its history.»

<sup>47</sup> Para uma definição de função própria cf. Millikan 1984:28.

são de algum modo mediadas pelo presente [2001:461]. Porém, se a teoria etiológica é histórica no sentido em que não é fundada em causas actuais (apesar de ser designada de “etiológica”), então não há nada que garanta que funções historicamente definidas sejam ainda causalmente eficazes no presente, o que faz dela causalmente *epifenoménica* [na terminologia de Bickhard 2009:557].

Outro problema da teoria etiológica reside na adopção da tese de que, a cada função, se encontra associada uma determinada explicação histórica de como a função foi desempenhada [Millikan 1993:243], explicação essa sem a qual não é possível determinar qual é exactamente a função de cada traço, pelo que atribuições de funções dependem inteiramente da nossa capacidade de conhecer qual foi a história natural de cada um desses traços [Millikan 1984:28] (o que pode ser muito difícil de se saber, senão por vezes impossível, tendo em conta exemplos paleológicos). Isto é bastante problemático porque requer que nos abstenhamos de atribuir funções a inúmeras características dos seres vivos somente porque não conhecemos a sua evolução – o que parece inteiramente contraproducente. Sober dá conta dum caso histórico famoso: ainda no séc. XVII, William Harvey, no seu *De Motu Cordis* [1628], atribuiu correctamente ao coração a função de bombear o sangue quando se apercebeu da estrutura anatómica das válvulas e apesar de não conhecer a sua história evolutiva [cf. Sober 1993b:85]. Para nós será de maior relevância outro exemplo: não é inteiramente claro qual foi a história natural dos nossos mecanismos cognitivos, existindo inúmeras teorias concorrentes acerca de como é que o encéfalo e todos os mecanismos que o compõem evoluíram do modo como evoluíram. Sendo este o caso, de acordo com a teoria etiológica, encontramos-nos impossibilitados de atribuir funções aos nossos mecanismos cognitivos enquanto não houver concordância acerca da sua história e isto apesar de estes parecerem ser tão úteis em tantas situações do nosso dia-a-dia.

Piorando a situação epistémica da teoria etiológica, não é só qual história é a correcta que importa. A própria questão relativa à existência ou inexistência de uma história torna-se relevante. Para exemplificar este ponto, os arguentes da teoria etiológica importaram uma experiência de pensamento de Donald Davidson: o “homem-do-pântano” (*swampman*) [Davidson 1987:443-4]. A experiência, como descrita pelo autor, envolve o próprio Davidson a andar por um pântano e a ser atingido por um relâmpago. Com o relâmpago, Davidson dá lugar a outro indivíduo que é idêntico a Davidson com a única excepção de que começou a existir naquele momento. Isto implica que Davidson e o homem-do-pântano são causalmente indiscerníveis, pelo que têm os mesmos valores de fitness. Contudo, Davidson teve uma história evolutiva, enquanto o homem-do-pântano não tem uma história evolutiva propriamente dita, pelo que, de acordo com a teoria etiológica, o homem-do-pântano não tem funções próprias, e isto apesar de ser tão fit quanto Davidson.

Autores como David Papineau pensam que o homem-do-pântano é um pseudo-problema [1987:75, ou 2001:279], mas nós acreditamos que esta experiência de pensamento realça um aspecto da teoria com bastantes consequências indesejáveis (apesar de não a refutar). Além do mais, não são sequer necessários exemplos fictícios para fazer jus a este problema. Pensemos em qualquer *exaptação*. Exaptações, ou, na terminologia com uma semântica teleológica de Darwin, *preadaptações*, ocorrem quando, na ausência de alteração estrutural, uma característica anterior adquire uma nova função biológica, quer por alterações ambientais, quer por alterações noutras características do próprio detentor dessa estrutura, incluindo mudanças comportamentais [Gould & Vrba 1982: em especial p.6; cf. ainda Ardila 2016:1]. Outros exemplos há que são relevantes para a nossa teoria cognitiva. Bickhard [2009:556-7], Jamie T. Whyte [1993:59-60], e Marcin Miłkowski [2015:2-3] sugerem que pensemos um exemplo em que, por via de mutações, surge (e.g.) um novo mecanismo cognitivo – até ao momento sem história selectiva – com capacidade de produzir crenças nunca antes instanciadas e que sucede ser bastante vantajoso na medida em que maximiza substancialmente a fitness do seu portador. Face



a estes exemplos, parece que a teoria etiológica não consegue dar conta de traços com utilidade corrente, nem de traços cuja utilidade é agora diferente daquela que lhes permitiu ser inicialmente seleccionados devido a (e.g.) mudanças ambientais [cf. ainda Gould & Lewontin 1979:252], o que levou alguns autores a considerar a teoria como “adaptacionista” [e.g. Sober 1993b:84].

Contudo, a teoria etiológica permite salvaguardar alguns dos usos comuns do termo *função* em Biologia e medicina, pelo que mesmo alguns dos seus oponentes reconhecem-lhe vantagens [e.g. Christensen & Bickhard 2002:4]. Considere-se o seguinte exemplo de Karen Neander [1991]: dum coração regularmente disposto a taquicardia diz-se, ainda assim, que tem a função de bombear o sangue, e mesmo que o órgão cesse por inteiro de bombear o sangue pelo corpo (e.g.) devido a um enfarte do miocárdio, não se deixa de dizer do coração que a sua função é bombear o sangue. Ao fim e ao cabo, o coração foi seleccionado *por* desempenhar esse papel no biosistema.

Passamos a dar início à exposição da teoria interactivista de funções biológicas e de como é que ela responde a estes desafios colocados à teoria etiológica. Começemos por nos lembrar de algumas noções fundamentais da teoria interactivista. Vimos que é da sincronização das dinâmicas metabólicas e das interacções com o ambiente que depende a recursividade da auto-manutenção dos biosistemas. De acordo com os interactivistas, isto é tudo quanto basta para que possamos extrair uma noção de normatividade natural. Isto é, eles crêem que é possível efectuar uma avaliação da capacidade do sistema ter sucesso ou falhar em persistir como um indivíduo autónomo e estabelecer normas que assim o determinem.

Importantemente, com isto não se quer dizer que a autonomia tem valor intrínseco e que estabelece normas compulsórias; pelo contrário, a autonomia ou estabilidade fora do equilíbrio termodinâmico não tem valor por si, sendo não obstante uma condição necessária à vida. Ou seja, a teoria interactivista diz-nos que *a vida que permanece viva é a que respeita as normas autonómicas*, mas nada nos diz relativamente a se a vida *deve* permanecer viva<sup>48</sup>, i.e. se as normas devem ser respeitadas.

Enquanto *H. sapiens* é possível reconhecer as normas autonómicas dos biosistemas como elas se apresentam no nosso quotidiano: sob a forma doutras normas que delas se derivam, de que são exemplo inúmeros desejos como “come”, “descansa”, “reproduz-te”, etc., mas que, em última instância, se podem perceber na terminologia interactivista, “auto-regenera-te”, “auto-regula-te”, etc.. São estas normas que determinam o comportamento dos biosistemas<sup>49</sup>. De acordo com a teoria interactivista, tudo quanto satisfaz estas normas, preservando em última instância a recursividade da auto-manutenção dos biosistemas, diz-se funcional. Ou seja, *se algo serve uma função, então está actualmente a promover a recursividade de auto-manutenção dum sistema* [cf. e.g. Bickhard 2001:462 e 2009:555]. Enquanto tal, podemos entender a autonomia como uma «restrição dinâmica e normativa básica à organização funcional dos sistemas adaptativos»<sup>50</sup> [Christensen & Bickhard 2002:3] e a história da evolução da autonomia como a história da evolução das funções biológicas.

---

<sup>48</sup> Os interactivistas creem que assim dissolvem o problema humeano da naturalização da normatividade, ultrapassando a chamada “falácia naturalista”.

<sup>49</sup> Para autores como Moreno e Mossió, p.ex., estas normas afectam causalmente os processos autonómicos sem serem afectadas pelos processos em si. Contudo, nós seguimos Hooker [2015] e Jason Winning & William Bechtel [2016:448], consideramos esta transcendência problemática, e adoptando ao invés uma concepção imanentista das normas.

<sup>50</sup> «a basic normative dynamical constraint on the functional organization of adaptive systems»

Normas que estabelecem o que é funcional no sentido de preservação da autonomia dizem respeito a condições *actuais* dos biossistemas e, enquanto tal, não são históricas. Note-se ainda que estas normas dizem respeito não só às condições interiores ao sistema, mas também ao arranjo das propriedades no meio ambiente, o que nos permite compreender que determinadas características do meio e exteriores ao biosistema sejam funcionais (um exemplo óbvio será o de um objecto construído por um *H. sapiens* com o fito último de preservar a sua própria autonomia). Em suma, porquanto o biosistema e o ambiente têm propriedades finitas e definidas, as normas determinam dentro do conjunto de possíveis interacções<sup>51</sup> dos biossistemas quais é que são funcionais.

Face aos desafios da teoria etiológica repare-se que, na medida em que o que é funcional preserva actualmente a autonomia, tanto o exemplo da função do coração descoberta por Harvey, como o de Davidson, o homem-do-pântano, como as exaptações, e os mecanismos cognitivos nunca antes instanciados, são salvos: de acordo com a teoria interactivista todos eles têm funções, que são também aptações. Porém, o que poderá o interactivista dizer dum órgão disfuncional? A teoria etiológica facilmente dá conta da intuição de que mesmo um coração que sofreu um enfarte tem a função de bombear o sangue. A solução dos interactivistas é engenhosa: a função esperada de um coração disfuncional pode ser derivada a partir da noção de função interactivista, bastando para isso que se pense que *alguns traços servem funções dependendo de outras funções serem servidas*, isto é, há funções de primeira e de segunda ordem<sup>52</sup> [Bickhard 2009:558-9]. Neste sentido, é possível dizer que o coração disfuncional *tem uma função que não está a ser servida*, permanecendo também possível dizer que há traços que podem servir funções sem por isso terem funções (no sentido em que não há outros traços que deles dependem para servirem funções). Ao fim e ao cabo, só dizemos que o coração que não funciona bem desempenha ainda assim uma função porque sabemos que outras funções dependem dele (e.g. a função das artérias e das veias, dos pulmões, fígado, rins, intestino, etc., da respiração e da nutrição e seu contributo último para a manutenção da autonomia do biosistema).

Ora, dada a longa história evolutiva dos actuais seres vivos, será de esperar que muitas das características dos biossistemas sirvam funções de segunda ordem. É inclusive de esperar que a evolução tenha produzido muitos traços de ordens inferiores que se encontrem *entrincheirados* (*entrenched*) [cf. Wimsatt 1997], querendo com isto dizer que, do seu funcionamento, depende o serviço de um maior número de funções e que o seu mau funcionamento acarreta inúmeras outras consequências para a autonomia do sistema como um todo (sendo que o entrincheiramento aumenta consideravelmente de ordem para ordem; e.g. o enfarte do miocárdio poderá ter consequências muito mais drásticas para o *H. sapiens* quando comparado com uma câibra do quadríceps femural ou mesmo uma apendicite aguda que nos obrigue à remoção do apêndice cecal). Devido ao nível de constrangimentos que o entrincheiramento impõe sobre a manutenção da autonomia, William Wimsatt prevê que as características entrincheiradas deverão persistir inalteradas durante mais tempo ao longo da história evolucionária [cf. Wimsatt 2007a:101].

---

<sup>51</sup> Um termo que nos permite referir a estas possíveis interacções é *affordances* [Gibson 1979:127].

<sup>52</sup> Também chamadas “*alopoiéticas*” [cf. Maturana & Varela 1972:68 & 135].

## 4. A Natureza e Função Mínima da Cognição

Até agora salientámos duas características fundamentais da natureza dos biosistemas: eles são autónomos e evoluem. Neste capítulo investigaremos brevemente outra característica dos biosistemas designada *cognição*.

A literatura em torno da definição de cognição é bastante extensa e encontra-se sobremodo preocupada com os limites desta capacidade. Num extremo do debate encontramos os defensores de uma perspectiva inflacionária que equaciona a vida com a cognição [e.g. Stewart 1996], noutro encontramos os defensores da perspectiva antropocêntrica e conservadora que equaciona a cognição com as faculdades ditas “superiores” dos sistemas cognitivos centralizados típicos da espécie *H. sapiens*. Estas encontrar-se-iam dedicadas a processos de inferência e *decisão racional*, entre outros. Neste ensaio iremos optar por uma posição que atribui cognição a todos os organismos com *órgãos especializados para o processamento da informação*<sup>53</sup> *advinda do ambiente, não necessariamente centralizados, mas ainda assim neuronais, e que têm em vista o controlo do comportamento flexível*<sup>54</sup>.

O nosso ponto de partida é novamente a discussão em torno dos modelos da bactéria com quimiotaxia apresentados no §1.1:10-1 [cf. novamente Bich & Moreno 2016:15]. No primeiro modelo vimos que não há processamento de informação do ambiente externo ao biosistema, somente reacção às condições interiores ao sistema (reacções estas que, em todo o caso, também mudam de acordo com as relações espaciais do sistema face ao ambiente). Neste caso, consideramos que não há cognição. No segundo caso, as reacções aos *inputs* detectados no meio são acopladas e inteiramente determinadas pelo actual estado estequiométrico do sistema, independentemente da existência de um órgão exclusivamente dedicado ao seu processamento. Também neste caso consideramos que não há cognição, ainda que haja processamento de informação proveniente do ambiente. No terceiro caso, o *input* é também exterior ao sistema, mas é já processado num órgão dedicado e que se encontra pelo menos parcialmente desacoplado da actividade metabólica do sistema. Este é um modelo *mínimo* de cognição.

Equacionamos o último modelo a uma definição minimalista de cognição porque só exibindo um mecanismo semelhante<sup>55</sup> ao do último modelo é que o organismo exibe uma das características típicas de organismos capazes de cognição: comportamento flexível regulado. Nos casos de amortecimento absorvente, característico do primeiro modelo, e da mera mediação metabólica, característica do segundo, não se verificam diferentes possibilidades de interacção entre organismo e ambiente sobre as quais se possa exercer controlo. Ainda que nestes casos o comportamento do sistema possa ser não-linear, continua dependente da detecção imediata de pistas ambientais. Concordamos assim com Godfrey-Smith [2002:241], para quem não há comportamento quando uma pista do ambiente determina *directamente* ou linearmente a actividade do organismo. É só quando há várias possibilidades de acção face à percepção imediata de uma pista do ambiente que se diz que há *comportamento* ou *agência adaptativa* [Moreno & Mossio 2015:§4.3]. Os primeiros dois modelos dão conta de respostas *fisiológicas* (seguindo a termino-

---

<sup>53</sup> A definição do conceito de informação, e em particular do conceito de informação biológica, tem sido alvo de muitas discussões, para as quais remetemos o leitor: Godfrey-Smith & Sterelny [2016]. Acreditamos que para os efeitos do presente ensaio não é necessário comprometermo-nos com uma definição para lá daquela do senso-comum.

<sup>54</sup> Um parêntesis é merecido para dizer que nem todos os interactivistas estão em acordo com esta teoria [cf. Moreno & Mossio 2015:167-8, n1].

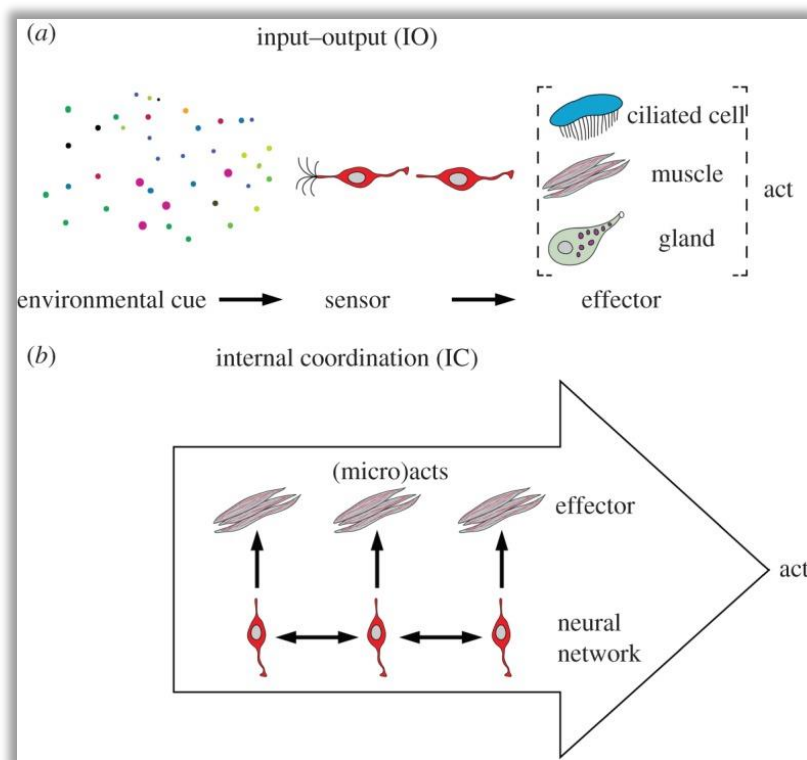
<sup>55</sup> Como *definição operacional* optámos por adoptar a terminologia de Ronald Giere segundo o qual a relação entre modelos e o mundo por eles descrito é uma relação de *semelhança* [cf. 1988:92-4 ou 2006].

logia de Rui Diogo [2017:1-6]), típicas de quando se verifica um estrito emparelhamento estequiométrico. É com o surgimento de um sistema de controlo desacoplado da actividade metabólica que surgem diferentes possibilidades de acção para o sistema como um todo.

Porém, não é somente o comportamento regulado que emerge com a presença de órgãos dedicados ao controlo. Os modelos não são iguais quantitativamente. Considerando o primeiro caso, qualquer interacção iniciada por resposta a alterações ambientais é dependente unicamente da velocidade do ritmo metabólico, enquanto, no segundo caso, esta velocidade poderá ser acrescida porque há detecção directa de sinais, apesar de a resposta continuar dependente do ritmo metabólico. É só no terceiro caso que a resposta iniciada é independente do metabolismo. Os três modelos exibem diferenças qualitativas e quantitativas, respondem diferentemente às pressões do meio, e podem ser diferencialmente seleccionados. A natureza da cognição ganha então especial relevância quando considerada sob a perspectiva autonómica e evolutiva.

Tendo recapitulado estes modelos perguntamo-nos agora qual a função da cognição. São várias as teorias acerca da função da cognição; de acordo com a nossa definição interactivista e evolutiva de função, sabemos *a fortiori* que a função última da cognição será preservar a autonomia do cognoscente. Contudo, isto não nos basta para diferenciar o papel específico que a cognição desempenha nas interacções com o ambiente, nem nas dinâmicas interiores do sistema, entre as quais o papel que lhe é pressuposto por outras características funcionais do sistema, assumindo algum nível de entrincheiramento.

Uma súpula de recentes estudos acerca da evolução dos sistemas nervosos ou cognitivos avançou a hipótese de que sistemas nervosos apareceram primeiramente com a função de *servir o controlo efectivo dos seus portadores em termos de coordenação interior e em termos de relações de input-output com o ambiente*, conforme apresentado pelo seguinte esquema [Jékely et al. 2015:2]:



Dois modelos para a origem evolutiva e função dos sistemas neuronais.

Nestes dois modelos de coordenação interior e de relações com o ambiente, o modelo (a) assemelha-se ao primeiro, e o modelo (b) ao segundo e terceiro modelo de cognição descrito por Bich & Moreno [2016], respectivamente. Estes estudos, a que podemos equiparar outros [e.g. Sterelny 2003; van Gelder 2005:176; Stout 2010, e Wolpert 2011], concordam que a função da cognição é controlar o comportamento dos organismos de forma a que a resposta sistémica aos estímulos que os organismos recebem do ambiente seja adaptativa, o que, no quadro da teoria interactivista, se deixa subsumir em última instância sob a função de preservar a autonomia<sup>56</sup>. Estes autores acreditam ainda que qualquer outra função que se possa atribuir à cognição se reduz inadvertidamente à função mais entrincheirada de exercer controlo adaptativo sobre o comportamento<sup>57</sup>. Peter Godfrey-Smith cunhou esta tese acerca da função da cognição sob a designação *Tese da Complexidade Ambiental*, que enunciou da seguinte forma:

«A função da cognição é permitir ao agente lidar com a complexidade ambiental.» [1998(1996):3, e novamente em 2002:236]

Assim formulada, a tese pressupõe que a função da cognição é controlar o comportamento com o objectivo de lidar com a complexidade ambiental<sup>58</sup> [cf. Godfrey-Smith 2002:238]. Esta tese parece ser defensável na medida em que alguma flexibilidade comportamental parece ser necessária para o sucesso evolucionário quando os ambientes são complexos, e isto, como nos diz Kim Sterelny [2003:11], porque comportamentos inflexíveis parecem não ter tanta utilidade. A bactéria com comportamentos quimiotácticos permite-nos ilustrar esta tese. Se o ambiente da bactéria for pouco complexo no que toca à *informação* acessível à bactéria (e.g. um ambiente homogéneo relativamente à concentração de glucose), o comportamento da bactéria é pouco substancial na determinação da sua fitness e não há quaisquer benefícios adicionados aos custos da produção e manutenção de um sistema cognitivo que controle este comportamento. Porém, se o ambiente da bactéria for muito informativo, então o número de possíveis interacções da bactéria com o seu meio que preservam ou não a sua viabilidade aumentam, sendo vantajoso regulá-las.

Estes autores concluem então que há uma *tendência macroevolucionária (macroevolutionary trend)* tal que, quanto mais complexo for o ambiente, mais complexo será o comportamento *apto*, havendo selecção *positiva* dos comportamentos mais complexos, por sua vez controlados por sistemas cognitivos também mais complexos, o que fortuitamente se traduz numa dinâmica causal de *feedback positivo* segundo a qual maior complexidade promove a evolução de maior complexidade<sup>59</sup> [Moreno & Mossio 2015:109 parecem concordar com esta tese macroevolucionária]. Isto, por sua vez, confere apoio à afirmação de Hooker segundo a qual a evolução da cognição pode ser pensada como «os estágios sucessivos de organização regulatória

---

<sup>56</sup> Pace Lewontin [1998] e outros cépticos relativamente à possibilidade de estudar a evolução *selectiva* da cognição em geral.

<sup>57</sup> A título de exemplo considere-se a *hipótese da inteligência social*, que considera que a cognição tipicamente humana se deve a uma resposta adaptativa dos agentes lidarem com complexidade social [cf. Humphrey 1976].

<sup>58</sup> Para simplificar, por *complexidade* queremos aqui dizer *diferença, heterogeneidade, e existência de múltiplas partes e propriedades*. Para uma discussão mais pormenorizada do termo cf. Wimsatt 2007b.

<sup>59</sup> Para uma exploração mais aprofundada deste tema: Sober [1997] oferece uma crítica da tese da complexidade ambiental; Sterelny [2003] oferece desenvolvimentos da tese de Godfrey-Smith [1998(1996)]; e Christensen [2010] oferece uma crítica de Sterelny [2003].

da interação com referência à autonomia, de ordem superior, centralizada neural, até que as capacidades humanas contemporâneas sejam alcançadas.»<sup>60</sup> [Hooker 2009:516]<sup>61</sup>.

---

<sup>60</sup> «the successive stages of autonomy-referenced, higher order, centrally neural, regulatory organisation of interaction, until contemporary human capacities are reached»

<sup>61</sup> Interpretamos Hooker como nos dizendo que a cognição humana é especialmente hipertrofiada comparativamente às restantes. Existem, não obstante, outras formas complexas de cognição não-humana potencialmente tão ou mais eficientes que a humana [cf. e.g. Godfrey-Smith 2016 e de Waal 2016].

## II. FORMULAÇÃO DA TESE DO TROPISMO AUTONÓMICO PARA A VERDADE

Nesta secção são apresentadas as principais teses deste ensaio, tendo por base as teorias expostas na secção anterior. Iremos abordar a natureza da relação entre sucesso interactivo e a verdade, argumentando a favor da existência de uma relação positiva entre a detenção de representações verdadeiras acerca do Mundo e a exibição de sucesso interactivo e evolutivo, à qual chamaremos *tropismo autonómico para a verdade*.

### 1. Antecipações como Portadores Primitivos de Conteúdo Pré-cognitivo

«E uma das funções mais gerais dos organismos vivos, uma das mais profundas, é olhar em frente, produzir futuro, como disse Paul Valéry.»<sup>62</sup>  
[Jacob 1982:66]

«O homem pode incorporar a verdade, mas não a pode conhecer.»<sup>63</sup>  
[Yeats 1939]

Em acordo com a natureza interactiva dos biosistemas, o exercício de uma função está dependente de certas condições estarem presentes no interior do organismo e no seu ambiente circundante, que satisfaçam a preservação da autonomia. É da maior relevância observar que as *condições de satisfação da autonomia* se encontram no *futuro* das interacções iniciadas, e isto não só pelo simples motivo que as interacções demoram tempo, mas também porque a satisfação da autonomia pode depender de condições ambientais ou organísmicas *potenciais*, i.e. *apenas actualizáveis após e devido às próprias interacções*. Assim, qualquer disposição de um sistema autónomo é tal que, ou *antecipa*<sup>64</sup> *condições favoráveis à manutenção da sua autonomia*, ou então não é viável para o seu portador, contribuindo negativamente para a sua fitness.

Bickhard chama a estas condições favoráveis à manutenção da autonomia “*pressuposições funcionais dinâmicas*” (*dynamic functional presuppositions*) [cf. e.g. Bickhard 2001:462], e elas dizem respeito a um estado, evento ou processo futuro<sup>65</sup> [cf. Christensen & Hooker 2000:136 & Hooker 2009] do Mundo. Sendo as antecipações *acerca* de um estado do Mundo, elas são *intencionais*<sup>66</sup> [cf. Jacob 2014], ainda que não necessariamente mentais (contra Brentano). Porém, o que é mais importante reter aqui é que isto significa também que as disposições dos sistemas autónomos *têm conteúdos cujas condições de identidade são dadas pelas condições de sucesso ou de satisfação da autonomia*. Respeitando o princípio semântico de bivalência, as *antecipa-*

---

<sup>62</sup> «And one of the deepest, one of the most general functions of living organisms is to look ahead, to produce future as Paul Valéry put it.»

<sup>63</sup> «Man can embody truth but he cannot know it.»

<sup>64</sup> Ao longo deste ensaio utilizaremos de igual modo *previsão*, *antecipação* e *expectativa* para nos referirmos a este fenómeno de antevisão do futuro.

<sup>65</sup> Vamos simplificar a discussão referindo-nos sempre ao *estado de coisas do Mundo*.

<sup>66</sup> Ainda que não-mentais (não necessariamente, pelo menos) no sentido comum de Filosofia da Mente e da Linguagem.

ções dos sistemas autónomos são verdadeiras quando são funcionais, e falsas quando são disfuncionais. Isto é, se as condições de satisfação da autonomia não se verificam, então os conteúdos das representações do próprio organismo são falsificados.

Com Bickhard [e.g. 2009:570], propomos que isto é quanto basta para que possamos obter uma *teoria minimalista da representação*, segundo a qual a representação é um fenómeno natural característico de todos os biosistemas e nos quais todas as disposições de organismos podem ser pensadas como *veículos representacionais* com conteúdos determinados. A representação é então um fenómeno abundante na Terra, entre os sistemas autónomos [cf. Papineau 2003].

Ainda que se distancie dos programas pragmatistas originais de autores como William James, a teoria interactivista da representação aqui traçada é *pragmatista* devido ao papel *fundamental* nela desempenhado pelas interacções dos biosistemas [Bickhard 2001:464 e 2009:547]. É também uma teoria naturalista ou reducionista – uma *Biosemântica*, na terminologia de Millikan [1989] – porque nos providencia uma *explicação do conteúdo de representações em termos das suas características causais ou funcionais, abandonando os termos intencionais próprios do discurso não-reducionista* [cf. Shapiro 1997:309-11]; no caso concreto, *dando conta da instanci-ção de fenómenos intencionais através de propriedades não-semânticas* [Loewer 1997:108] (e.g. as propriedades termodinâmicas dos sistemas autónomos).

Às condições de satisfação da autonomia pressupostas pelas interacções chamaremos “conteúdos primitivos” porque os mesmos *não são proposicionais* no sentido *linguístico* habitual de *proposição* [sensu Hooker 1975, Christensen & Hooker 2000:140, Bermúdez 2003, e Bermúdez & Cahen 2015]<sup>67</sup>. De facto, conteúdos primitivos são usualmente *implícitos* e não explícitos, podendo inclusive ser *somáticos* ou incorporados (*embodied*) e *não mentais*, nem linguísticos. Porém, se identificarmos proposições com portadores de verdade, então a teoria interactivista da representação é proposicional (deixamos ao juízo do leitor decidir se disposições somáticas têm ou não conteúdos proposicionais). Não obstante esta problemática sobremodo conceptual, a teoria interactivista propõe um modelo tripartido de representação semelhante aos que se encontram na literatura e de acordo com o qual existem [cf. Bickhard 2003:103]: (1) representações; (2) conteúdos; e (3) aquilo acerca de que versam os conteúdos. Mas no modelo interactivista estes elementos podem ocorrer sob outra terminologia, nomeadamente: (1) interacções antecipatórias; (2) as pressuposições funcionais dinâmicas; e, finalmente, sem terminologia especial, (3) os objectos do mundo, respectivamente.

Por outro lado, porque a teoria se compromete, contra a ortodoxia, com a existência de conteúdos pré-cognitivos, ela aproxima-se de algumas propostas que consideram *adaptações* [e.g. adaptações de um olho a fotões, Vollmer 1975] e alguns *sistemas hereditários* [e.g. o genoma, Shea 2007] como sendo *representantes*; ainda que também divirja destas propostas, na medida em que estas dizem respeito a conteúdos estabelecidos «filogeneticamente *a posteriori*»<sup>68</sup> – para empregar uma expressão de Konrad Lorenz [1941] –<sup>69</sup>, enquanto no nosso caso

---

<sup>67</sup> Iremos suspender o juízo relativamente à possibilidade de expressar por princípio todos e quaisquer conteúdos primitivos por frases, pois não é claro para nós que indexicais, mapas cognitivos, perceptos, entre outros, possam receber tal tratamento.

<sup>68</sup> O que significa que a *covariância* entre representações e aquilo que estas representam se dá à velocidade da evolução por selecção natural, de acordo com o modelo etiológico.

<sup>69</sup> Atente-se ao seguinte excerto de Nicholas Shea: «genes’ representational contents are fixed by their selectional properties, and so can figure only in phylogenetic explanations.» [Shea 2007:327]. Aqui não atenderemos às diferenças entre a nossa teoria e a teoria de Shea, que se funda em elementos de uma teoria etiológica de funções biológicas. Contudo, concordamos com a resposta de Shea aos críticos da sua



avaliamos o seu valor de verdade independentemente da existência de uma história evolutiva e relativamente à sua contribuição actual para a autonomia (*i.e* enquanto *aptações*).

A título de exemplo considere-se um uso corrente de *representação* em Biologia na esteira da definição semântica de fitness ecológica. Dizem-nos os biólogos que uma determinada adaptação representa uma resposta a condições ambientais específicas. Peter Richerson & Robert Boyd, p.ex., dizem-nos que as «tendências Cenozóicas de complexidade cognitiva *representam* adaptações a um ambiente cada vez mais variável»<sup>70</sup> [2000:329]<sup>71</sup>. Mas qualquer adaptação seguirá este esquema, tendo por conteúdo as condições de selecção *positivas* sofridas pelo seu portador, pelo que facilmente se desenvolvem outros exemplos: (e.g.) um organismo aeróbico presente na Terra após O Grande Evento de Oxigenação (seja um qualquer dinossauro com pulmões) antecipa uma atmosfera rica em oxigénio e a não verificação duma atmosfera com tais qualidades e quantidades será disfuncional para este organismo.

Para além disto, compatibilizamos ainda a nossa abordagem com teorias *biosemióticas* [e.g. Barbieri 2008], para as quais o mundo se encontra repleto de significado até a um nível unicelular e molecular (enquanto concordarmos que estruturas deste tipo contribuem para a autonomia), ainda que este significado não seja explícito para os organismos. Como nos dizem Moreno & Mossió:

«Nesta visão o ambiente torna-se num mundo repleto de significado: factos que de fora podem parecer apenas puramente físicos ou químicos tornam-se em influências positivas, negativas ou neutras para o sistema, dependendo se contribuem, dificultam ou não têm efeitos na manutenção da sua identidade dinâmica. Mesmo o organismo mais simples cria um conjunto de partições preferenciais do mundo, convertendo interações com o seu meio circundante em valores elementares, (...) von Uexküll (1892/1940) chamou a este mundo subjectivo significante de cada organismo *Umwelt*.»<sup>72</sup> [2015:xxix]

Com a devida atenção a que estes *valores*, atribuídos a cada interacção, serão tácitos e intrínsecos no caso dos organismos sem mecanismos cognitivos ou mesmo com capacidades cognitivas rudimentares, porquanto a capacidade de atribuição de valores explícitos a interacções ou *potenciais de interacção* (na terminologia de Bickhard [cf. 2009:570-1]) parece ser *meta-cognitiva* [cf. Shea 2012]. Não obstante, e ainda que tacitamente, o *Umwelt* de cada organismo, o

---

teoria (e.g. Godfrey-Smith e Papineau), quando este nos diz que eles são críticos desta possibilidade porque *não pensam a variação destas representações relativamente ao ambiente numa escala temporal filogenética*. Somente nesta escala é que verificamos algum tipo de correlação entre características seleccionadas e condições ambientais, sendo que alterações ambientais podem conduzir a falhas explicáveis em termos de conteúdos, *i.e.* a falsidade dos conteúdos do genoma explicam o insucesso dos seus portadores [cf. Shea 2007:326].

<sup>70</sup> «[W]e argue that Cenozoic trends in cognitive complexity *represent* adaptations to an increasingly variable environment.», ênfase nossa.

<sup>71</sup> É de notar que estes autores concordam com a tese da complexidade ambiental [§1.4.], apesar de não a mencionarem pelos mesmos termos.

<sup>72</sup> «In this view, the environment becomes a world full of significance: facts that from the outside may appear just as purely physical or chemical develop into positive, negative, or neutral influences on the system, depending on whether they contribute to, hinder or have no effect on the maintenance of its dynamic identity. Even the simplest living organism creates a set of preferential partitions of the world, converting interactions with their surrounding media into elementary values, (...) von Uexküll (1892/1940) called this subjective meaningful world of each organism *Umwelt*.»

seu mundo fenoménico e de significado [cf. von Uexküll 1934 & 1982] é «povoado por *affordances*»<sup>73</sup> [Dennett 2017:79], por aquilo que «o ambiente oferece ao animal, o que providencia ou fornece, seja para o bem ou para o mal»<sup>74</sup> [Gibson 1979:127].

Será elucidativo observar que, de acordo com esta teoria semântica, os conteúdos das representações são determinados a par da contribuição das propriedades causais dos veículos de representação para a fitness, isto é, *relativamente a um ambiente*. A teoria é, enquanto tal, *externista* – o que não impede que sistemas cognitivos mais sofisticados tenham acesso às *razões naturais* pelas quais as suas intenções têm os conteúdos que têm. Porém, não é claro que um acesso alargado aos conteúdos seja sempre vantajoso para os organismos, dum ponto de vista evolucionário, devido a um princípio de economia da natureza. Neste contexto, o *slogan* de Daniel C. Dennett é esclarecedor: «competência sem compreensão é o caminho da Natureza»<sup>75</sup> [2017:84]. De facto, ter acesso às razões naturais pelas quais determinadas representações têm o conteúdo que têm é um problema complexo que, neste enquadramento, coincide com o problema epistémico da determinação da referência ambiental de um fit ecológico [cf. §1.2.2.]. Ou seja, não é *perceber* a razão natural pela qual  $x$  é fit em  $y$  que torna  $x$  fit em  $y$  (o que não quer dizer que perceber a razão natural não possa ser evolucionariamente vantajoso). Por outras palavras, ter uma determinada representação verdadeira é conducente ao sucesso independentemente de sabermos porque é que ela é conducente ao sucesso (e a nossa teoria procura explicar precisamente porque é que representações verdadeiras conduzem ao sucesso).

Continuando a qualificação da teoria interactivista, podemos dizer que é uma teoria de semântica *vero-condicional* (*truth-conditional semantics*) porque *define o conteúdo pelas suas condições de verdade* ou como *aptidão para a verdade* (*truth-aptness*). Neste caso, as condições de verdade de cada representação são as condições sob as quais cada antecipação satisfaz a autonomia do seu portador, sendo esta relação entre antecipações e os seus conteúdos *necessária* ou, na terminologia da metafísica relacional, *interna* [Orilia & Swoyer 2016; cf. Schaffer 2010 para maiores detalhes e variações em definições de relações internas]. Já relativamente ao conteúdo em si, podemos verificar que este só é conducente ao sucesso, *i.e.* só é verdadeiro, quando o ambiente for de uma determinada maneira, *i.e.* quando o *referente* ou *estado de coisas* que garante o sucesso se verifica, respeitando ainda certos constrangimentos organizacionais do biosistema. Assim, como as pressuposições funcionais dinâmicas das interações são *verdadeiras* se e somente se o ambiente reflectir o seu conteúdo, podemos dizer que a relação entre *conteúdos verdadeiros* e o ambiente é também uma relação interna, ainda que *extrínseca*<sup>76</sup> (o mesmo não se podendo dizer de conteúdos falsos). E, por *transitividade*, podemos também dizer que há uma relação interna entre antecipações verdadeiras e os poderes do mundo que estas referem e, aliás, com os quais interagem. Contudo, o estudo da *ecologia semântica* das representações e suas contribuições para a fitness ecológica encontra-se incompleto sem a contribuição das dinâmicas interiores do organismo. Por este motivo, a semântica é então organismo-dependente, *i.e.* a semântica é dependente da coerência interior do organismo. Estas mesmas dinâmicas, que anteriormente designámos de *holismo da fitness*, reflectem-se aqui sob a forma de *holismo semântico*.

---

<sup>73</sup> «an Umwelt populated by affordances». Optámos por não traduzir ‘affordances’ na ausência de uma expressão satisfatória que não seja a sua própria definição.

<sup>74</sup> «the environment offers the animal, what it provides or furnishes, either for good or ill»

<sup>75</sup> «Competence without comprehension is Nature’s way»

<sup>76</sup> Bickhard [2003:103] nega esta tese devido a uma crítica de Bertrand Russell ao idealismo dos filósofos britânicos Green e Bradley. Contudo, todos estes filósofos assumem que *relações internas são necessariamente simétricas*. Agradecemos a David Yates a sugestão de *rejeitar o compromisso com a simetria das relações internas*.

Vimos que a determinação do conteúdo primitivo é dada pelas condições de sucesso das interações na manutenção da autonomia. Mas o que dizer então da ecologia semântica de uma representação quando (e.g.) uma aptação ou adaptação a um determinado ambiente é igualmente vantajosa noutro ambiente? Uma primeira intuição poderá levar-nos a concluir que uma alteração do ambiente de referência deveria corresponder ou a uma alteração ou a uma falsificação do conteúdo. Mas isto não é correcto. As condições de sucesso na manutenção da autonomia são abstractas o suficiente para se generalizarem acerca de todos os ambientes em que dada representação contribui para a manutenção da autonomia. Ao fim e ao cabo, *as propriedades* das interações antecipatórias ou veículos representacionais *não mudam qualitativamente* com o ambiente (o que é compatível com a *emergência de qualidades* interactivas mediante novas relações com o ambiente, em virtude de relações extrínsecas [cf. Santos 2015a e 2015b])<sup>77</sup>.

Isto poderá ficar também mais claro se pensarmos que as condições de sucesso de determinada antecipação são poderes causais do ambiente, e se pensarmos que diferentes estruturas podem realizar o mesmo poder causal. Por consequência, é-nos requerido que identifiquemos as condições de sucesso como aquilo que se mantém *invariante* ou comum em todos esses diferentes ambientes (p.ex., um determinado poder causal que qualifica um tipo de objectos). A alteração do conteúdo em si só se dá com uma alteração física do veículo representacional (e.g. através da evolução por selecção natural ou com o surgimento de novas mutações). E a falsificação do conteúdo só se dá quando determinada representação não é *potencialmente* vantajosa. Interessantemente, esta proposta permite que determinadas disposições dos biosistemas sejam vantajosas (ou aparentemente vantajosas) sem que encontrem as suas condições de sucesso (e.g. porque o ambiente em que elas seriam vantajosas nunca chega a ser experienciado, mas existe). Porém, isto pode dar-se devido a algum tipo de acaso, mas nunca necessária ou sistematicamente. Uma disposição sempre conducente ao sucesso diz-se apta ao ambiente e encontra nele as condições de verdade que determinam o seu conteúdo.

## 2. Antecipações como Veículos de Representação Desacoplados do Ambiente

«Nós representamos para intervir,  
e intervimos para representar.»<sup>78</sup>  
[Hacking 1983:31]

Entendemos que a teoria minimalista da representação aqui traçada se pode estender a formas de representação mais complexas. De facto, entendemos que as representações mais complexas se *fundam* nas dinâmicas dos conteúdos primitivos, mesmo quando falamos de *veículos de conteúdos* tão sofisticados quanto *representações mentais*, preservando um aspecto fundamental da teoria interactivista da representação: *o conteúdo precede ontologicamente a cognição, emergindo sincronicamente com as condições de satisfação da autonomia*, tal qual o que sucede com a emergência de funções e da normatividade relativa à autonomia. Salvas algumas

---

<sup>77</sup> Para deixar mais claro, a diferença aqui é entre um objecto (no caso, uma interacção ou representação) adquirir ou perder qualidades, por um lado, e uma qualidade transformar-se noutra, por outro. A emergência diz respeito ao primeiro caso de surgimento de novas qualidades, enquanto o segundo no qual uma propriedade se transforma noutra nos parece ser metafisicamente impossível [cf. Santos 2015a].

<sup>78</sup> «We represent in order to intervene, and we intervene in order to represent.»

diferenças, cremos que o trabalho de Christensen nos permite dar os primeiros passos no sentido desta demonstração. Christensen [1996:315] propõe que pensemos as intenções e a sua *acerquidade* (*aboutness*) num *continuum* relativamente à sua complexidade, no extremo de maior complexidade (reconhecida) identificando os casos exemplares de *representações mentais* e no extremo oposto aquilo que aqui chamamos de conteúdos primitivos. Somos da opinião que os exemplos de complexidade a que Christensen se refere podem ser explicados inteiramente por mecanismos cognitivos, sendo que uma análise comparativa do modelo minimalista da cognição fundado em Bich & Moreno [2016] nos permitirá articular a teoria do *continuum* intencional com a teoria da cognição previamente apresentada [cf. I.4.].

Comecemos então por observar que todos os sistemas autónomos exploram inerentemente informação do seu ambiente [Wuketits 1990:53 e Gell-Mann 1994:348], variando no modo como o fazem (e.g. os três modelos de Bich & Moreno [2016]). Diz-nos Christensen [1996:316] que a intencionalidade varia em grau consoante esta exploração é realizada. As intenções mais simples têm aquilo a que temos vindo a chamar “conteúdos primitivos”, sendo que cada intenção se deixa qualificar por «aquilo que poderá ser chamado de objectivo “instintivo”: *i.e.* o ponto focal de um padrão de comportamento inflexivelmente sequenciado. Neste caso, pode não ser verdade que X intenciona Y como um estado de coisas no sentido em que X tem uma representação das propriedades funcionais relevantes de Y (ou mesmo qualquer representação de Y)»<sup>79</sup> [Christensen:1996:315], o que concorda com os comportamentos inflexíveis atribuíveis aos sistemas semelhantes ao primeiro e segundo modelo de Bich & Moreno [2016].

Porém, o desacoplamento estequiométrico de um órgão processador da informação contida no ambiente permite ao sistema exibir comportamentos distintos em resposta a uma mesma pista do ambiente. Isto é, um órgão de regulação de comportamento estequiométricamente desacoplado dá azo a um segundo tipo de desacoplamento relativo ao ambiente [cf. Sterelny 2003:30-1]. Isto poderá suceder por várias razões diferentes. P.ex., a resposta do sistema poderá depender de uma avaliação da fiabilidade da pista sensorial (se é um sinal ou ruído) relativamente a detecções anteriores, a curto ou longo prazo, do mesmo *tipo* de pista (e.g. detecções de fotões de comprimento de onda [x, y] que atingem a retina), não se deixando determinar unicamente pela pista actualmente detectada<sup>80</sup>. Enfim, esta plasticidade fenotípica do comportamento regulado pela cognição é susceptível a factores *ontogenéticos*, o que permite ao sistema responder adaptativamente a maior variabilidade ambiental durante o seu tempo de vida<sup>81</sup>. Esta adaptabilidade contrasta com a de organismos meramente capazes de comportamentos inflexíveis que, por sua vez, só podem ser seleccionados pela natureza através de um

---

<sup>79</sup> «[A] system may have what might be called an “instinctive” goal: *i.e.* the focal point of an inflexibly sequenced behaviour pattern. In this case it may not be true that X intends Y *as a state of affairs* in the sense that X has a representation of the relevant functional properties of Y (or any representation of Y at all).

<sup>80</sup> O modelo predictivo da cognição [cf. Friston & Stephan 2007; Friston 2010; e Clark 2015] procura formalizar esta possibilidade.

<sup>81</sup> N.b. sistemas de comportamento inflexível podem alterar as suas estratégias comportamentais durante a ontogénese devido (e.g.) à influência do ambiente em mecanismos epigenéticos. Contudo, esta alteração comportamental – que é ao fim e ao cabo um tipo de flexibilidade comportamental – é muito lenta, verificando-se ainda que de comportamento para comportamento a variabilidade é contínua e não discreta.

As representações mentais parecem ter o conteúdo parcialmente dependente de valores socio-culturais (com algumas excepções, e.g. a “ença” (*alief*) [e.g. Gendler 2008] de que cobras e aranhas são perigosas está mais-ou-menos distribuída por todos os *H. sapiens*) e, neste sentido, variam bastante de população para população, enquanto as representações pré-mentais parecem ser menos diversificadas dada a pouca plasticidade dos órgãos extra-encefálicos (salvo algumas afecções epigenéticas, que podem em todo o

moroso processo *filogenético* de *tentativa e erro*<sup>82</sup> – a expressão é de Donald T. Campbell [1974] –, o que poderá provar-se desvantajoso em condições de rápidas alterações ambientais.

O que é fulcral compreender é que o *armazenamento de informação acerca do ambiente no órgão cognitivo* (ou *memória cognitiva*) constitui em si mesmo uma estrutura do sistema que é alterável com a detecção de novos sinais (ou não haveria aprendizagem e consequente resposta adaptativa) e que *ceteris paribus* pode permanecer acessível ao sistema independentemente dessa detecção<sup>83</sup> (o que é também parcialmente responsável pela flexibilidade comportamental). Complicando este modelo, sistemas cognitivos mais complexos encontram-se munidos de mecanismos cognitivos que aumentam a flexibilidade da resposta comportamental por executarem processamento de informação *sobre* a memória cognitiva (e.g. por emprego de mecanismos de *inferência*). Estes sistemas mais complexos dizem-se então de capacidade epistémica superior [Christensen 1996:314]. Nas palavras de Murray Gell-Mann, um sistema adaptativo complexo é um sistema que «adquire informação acerca do seu ambiente e da sua interacção com o seu ambiente, identificando regularidades nessa informação, condensando essas regularidades num tipo de “esquema”, ou modelo, e agindo no mundo real com base nesse esquema»<sup>84</sup> [Gell-Mann 1994:348]. Gell-Mann parece assim concordar com Robert Rosen, para quem um sistema antecipatório é definido como «um sistema contendo um modelo predictivo de si próprio e/ou do seu ambiente, que lhe permite mudar de estado num instante de acordo com as previsões que o modelo faz com respeito a um instante mais tardio»<sup>85</sup> [Rosen 1985:339].

Em termos da teoria da representação, o que sucede é que estes sistemas adaptativos complexos ou antecipatórios passam a interagir com o mundo de um modo que é (pelo menos) *parcialmente independente* de cada captura de informação do mundo, e parcialmente dependente do estado cognitivo do agente (do *modelo* de que nos fala Gell-Mann ou Rosen) e das dinâmicas inerentes à sua estrutura (entre as quais as inferenciais). Sendo estes sistemas capazes de vários tipos de comportamentos relativos à mesma pista, consoante o seu estado cognitivo, verifica-se então que *a condição de sucesso da disposição cognitiva não é mais dada por um único comportamento inflexível* (como no caso de representações somáticas), *mas por todos os comportamentos possíveis que se seguem de determinado modelo/estado cognitivo*. Este modelo é, pois, ele próprio, uma representação (mais complexa e possivelmente intencional no sentido “forte” de Christensen, *i.e.* no outro extremo do *continuum*) das condições de sucesso da manutenção da autonomia.

O desacoplamento das representações contribui ainda para a profundidade do problema do ambiente de referência. É já complicado atribuir ambientes de referência a adaptações ou a aptações, mas o problema fica ainda maior quando diferentes tipos de comportamento são associáveis a uma mesma representação. E, o que é mais agravante, que o mesmo tipo de comportamento possa ser associado a diferentes representações. Há uma relação de muitos para

---

caso não afectar substancialmente o conteúdo de representações somáticas e comportamentais inflexíveis).

<sup>82</sup> Nesta medida podemos também entender todas as adaptações como uma resposta “indutiva” ao ambiente.

<sup>83</sup> Poderemos argumentar no sentido de que a capacidade de memória e a capacidade de acesso à memória são distintas e independentes. Esta questão não pesa nos argumentos subsequentes.

<sup>84</sup> «(...) acquires information about its environment and its own interaction with that environment, identifying regularities in that information, condensing those regularities into a kind of “schema”, or model, and acting in the real world on the basis of that schema.»

<sup>85</sup> «A system containing a predictive model of itself and/or its environment, which allows it to change state at an instant in accord with the model's predictions pertaining to a later instant.»

muitos: representações diferentes podem dar azo a um mesmo comportamento e comportamentos diferentes podem ter por origem a mesma representação<sup>86</sup>. O problema da *atribuição de conteúdos específicos* aos agentes também não desaparece aqui, o que em todo o caso não nos impede de seguir a prática de muitos etólogos e psicólogos que utilizam *comportamentos expectantes* como *indicadores para atribuição de crenças* a animais não-humanos e a crianças numa fase pré-linguística (e.g. durante o balbucio da fase de lalação [e.g. Southgate *et al.* 2007 e Scott & Baillargeon 2009])<sup>87</sup>. Algumas representações desacopladas, de que são exemplo representações mentais, podem ainda ser proposicionais e expressas, enquanto tal, linguisticamente. Todavia, isto não quer dizer que o problema da referência ambiental seja resolvido, apenas que o conteúdo destas representações se encontra *explicitamente mais determinado*.

### 3. A Tese do Tropismo Autónómico para a Verdade

Em jeito de resumo do que foi dito até agora propomos então a tese fundamental deste ensaio. Consagraremos as restantes secções deste ensaio à aplicação, explicitação, e defesa da mesma:

Um tropismo autónómico para a verdade é uma relação interna e extrínseca entre conteúdos verdadeiros e as condições de verdade, *qua* condições de sucesso, das interacções dos seus portadores com o ambiente ou consigo próprios, em virtude da sua autonomia.

---

<sup>86</sup> Esta disparidade pode provar-se vantajosa (e.g. é vantajoso em termos do contributo para a fitness de uma presa que o seu comportamento seja imprevisível da perspectiva dos predadores), mas acarretando custos (e.g. a cooperação entre indivíduos da mesma espécie poderá tornar-se dependente dum mecanismo cognitivo capaz de processar uma *teoria da mente*) [cf. Sterelny 2003].

<sup>87</sup> Mas há autores que não hesitam em atribuir conteúdos a crenças de determinados animais; Wuketits [1990:70], p.ex., sugere que um gato acredita que um cão é perigoso, dizendo-nos que a crença é de algum modo adquirida pré-conscientemente. Este exemplo poderá assim aproximar-se do medo humano a *Serpentes* ou *Arachnida*. Contudo, utilizar o comportamento como indicação de uma atitude é diferente de utilizar o comportamento como indicação de uma proposição.

### III. A APLICAÇÃO DA TESE DO TROPISMO AUTONÓMICO PARA A VERDADE À SEMÂNTICA

Nesta secção iremos pensar a naturalização de uma *semântica de sucesso* com base na tese do tropismo autonómico para a verdade.

#### 1. Uma Semântica de Sucesso ou Pragmatista

Semânticas de sucesso são teorias semânticas *pragmatistas* que identificam o *conteúdo* de uma dada representação com as *condições de utilidade* ou de *sucesso* de uma *acção que é realizada à luz dessa representação* e cujo propósito é definido como a *satisfação de determinados desejos* [cf. Whyte 1990, Dokic & Engel 2002 e 2005, Bermúdez 2003, e Blackburn 2005]. Porquanto «[c]renças verdadeiras são funções de desejos para acções que causam os pensadores a comportarem-se de modos que satisfazem os seus desejos»<sup>88</sup> [Bermúdez 2003:65], José Luis Bermúdez [2003:104] chama às semânticas de sucesso “*funcionalistas*”.

De um modo mais abrangente, podemos dizer que semânticas de sucesso, também chamadas de “*semânticas de acção*” [Blackburn 2005:22], são quaisquer teorias semânticas que tomem o sucesso de uma acção como critério fundamental para uma teoria acerca dos conteúdos e da intencionalidade. Enquanto tal, elas fazem parte do conjunto mais geral de *teorias da representação centradas na acção* [cf. Anderson & Rosenberg 2008, e Gładziejewski 2016], conjunto a que também pertence o Interactivismo. Num sentido lato, podemos ainda dizer que semânticas de sucesso são teorias que procuram *derivar condições de verdade de condições de sucesso* [Dokic & Engel 2002:46].

#### 2. O Princípio de Ramsey nas Fundações da Semântica de Sucesso

Os actuais defensores e críticos de semânticas de sucesso identificam as origens históricas da teoria moderna entre os pragmatistas estadunidenses (e.g. James, Charles Sanders Peirce, e John Dewey) e entre os pragmatistas de Cambridge (Russell, Frank Plumpton Ramsey, e Henry Wilkes Wright) [cf. Whyte 1990, Godfrey-Smith 1994, e Boncompagni 2014]. Mas terá sido Ramsey quem mais contribuiu para a formulação actual da teoria, apesar das actuais divergências relativamente à interpretação da sua obra<sup>89</sup>.

No seu ensaio de 1927, «Factos e Proposições», Ramsey apresentou, entre outros assuntos, as suas reflexões acerca das interacções entre uma galinha e uma lagarta potencialmente venenosa. Ramsey propôs que pensássemos na galinha que se abstém de comer lagartas do tipo x (eventualmente devido a experiências passadas desagradáveis com lagartas daquele tipo) como

---

<sup>88</sup> «true beliefs are functions from desires to actions that cause thinkers to behave in ways that will satisfy their desires».

<sup>89</sup> cf. Bermúdez 2003:§4 e Methven 2015:§6 para aprofundar as diferenças entre o projecto inicial de Ramsey e a Semântica de Sucesso como desenvolvida a partir da década de 90.

sendo portadora da crença “lagartas x são venenosas” (à semelhança das anteriormente consumidas, que eram indigestas). Independentemente do problema da atribuição de conteúdos<sup>90</sup>, a lição de Ramsey que se encontra na origem das teorias semânticas de sucesso é a seguinte:

«[A] relação entre o comportamento da galinha e os factores objectivos era tal que as acções seriam úteis se e somente se as lagartas fossem realmente venenosas. Assim, qualquer conjunto de acções para a utilidade das quais  $p$  é uma condição necessária e suficiente pode ser chamada uma crença em  $p$ , e como tal seria verdadeira se  $p$ , i.e. se elas são úteis.»<sup>91</sup> [Ramsey 1927:159]

Tendo por base este ensaio de Ramsey e o subsequente «Verdade e Probabilidade» [1929], Nils-Eric Sahlin, comentador de Ramsey, sugeriu representar as possíveis interacções da galinha com a lagarta mediante uma tabela-matriz de decisão [1990:82] semelhante à seguinte:

	Não-venenosa	Venenosa
Comer	Nutritivo	Problemas gástricos
Não-comer	Falha uma alimentação	Evita problemas gástricos

A galinha pode então *agir ou não agir* face a uma pista ambiental que é *verdadeira ou falsa* (neste caso, comer ou não comer uma lagarta que é ou não venenosa). Atendendo às dinâmicas representadas nesta matriz simplificada é-nos permitido concluir que a interacção da galinha com a lagarta é tal que a galinha que detecta pistas verídicas (ou sinais) que denunciam se a lagarta é venenosa ou não está em melhores condições de maximizar a utilidade esperada das suas interacções com a lagarta. Isto torna-se claro na medida em que, na ausência de uma crença verdadeira acerca da toxicidade da lagarta, a maximização da utilidade esperada depende unicamente de uma análise da relação custo-benefício da interacção [Ramsey 1929; cf. Sahlin 1990:71-2]. Então, *uma acção em acordo com uma crença verdadeira acerca de um estado de coisas do mundo é uma acção que sistematicamente maximiza a utilidade esperada*.

É tendo por base estas observações e interpretação forte do exemplo de Ramsey que Jamie T. Whyte [1990] propõe o seguinte princípio, cunhado em nome de Ramsey:

<sup>90</sup> Atendendo à natureza relacional da tese do tropismo autonómico para a verdade, note-se que, quando se fala das propriedades causais de uma crença como *independentes* do seu conteúdo, isto é unicamente o resultado de se manter o pressuposto metafísico de que não há uma relação *interna* entre representações e os seus conteúdos. Dado que a tese do tropismo autonómico para a verdade pressupõe tal relação interna, mesmo não estando em melhores condições de atribuir conteúdos às representações mentais (e.g.) de uma galinha, podemos não obstante seguramente argumentar que aquela crença específica da galinha tem um e somente um conteúdo. Cf. Bermúdez 2003:§4-5 para mais desenvolvimentos relacionados com a aplicação de uma semântica de sucesso a comportamentos não-linguísticos (Bermúdez conclui com péssimismo que uma semântica de sucesso não nos permite distinguir quais os conteúdos da representação associada a um determinado comportamento [2003:64]). Como nota Blackburn [2005:22], Ramsey não desenvolveu estas questões. Methven [2015] oferece uma explicitação dos derradeiros compromissos de Ramsey.

<sup>91</sup> «the relation between the chicken's behaviour and the objective factors was that the actions were such as to be useful if, and only if, the caterpillars were actually poisonous. Thus any set of actions for whose utility  $p$  is a necessary and sufficient condition might be called a belief that  $p$ , and so would be true if  $p$ , i.e. if they are useful.»



*Princípio de Ramsey* <sup>Whyte</sup>: «A condição de verdade de uma crença é aquilo que garante a satisfação de qualquer desejo pela acção que aquela crença e desejo combinados causariam»<sup>92</sup> [Whyte 1990:150].

Não é necessário focarmos a nossa atenção na galinha para perceber este princípio. Qualquer crença de senso-comum em princípio deverá bastar. P.ex., se tivermos o desejo de beber água e acreditarmos que *k* “girando o manípulo, a torneira jorra água potável”, e obtivermos sucesso na acção de satisfazer a nossa sede é porque *k* é verdadeira. Por consequência, se aceitarmos a teoria semântica vero-condicional, o sentido de *k* será as condições nas quais obtivemos sucesso na nossa acção de obter água.

Contudo, se há crenças de senso-comum que nos permitem facilmente fazer sentido do princípio de Ramsey, a versão Whyteana do princípio também se depara rapidamente com exemplos que fazem dela contra-intuitiva. Imagine-se um agente, chamemos-lhe Pascal, que, após se servir de um café, pensa para si próprio, “Se despejar uma colher deste *açúcar* (que se encontra na taça ao lado da máquina do café) no meu café, este ficará adocicado, tal como gosto”. A acção de Pascal tem sucesso: ele age de tal forma que ao misturar uma colher daquele pó branco na sua chávena de café este fica mais doce. Tudo parece concordar com o princípio de Ramsey <sup>Whyte</sup>: Pascal teve sucesso na acção que conduziu de acordo com a crença *p* “Se despejar uma colher (...) este ficará adocicado” na medida em que conseguiu satisfazer o seu desejo. Contudo, achamos valor no velho cliché: as aparências iludem. Uma química entra na mesma sala e, após uma breve conversa entre os dois, Pascal descobre que ele não colocou açúcar (e.g. sucrose, C<sub>7</sub>H<sub>5</sub>NO<sub>3</sub>S) no seu café, mas sim um açúcar artificial (e.g. sacarina, C<sub>12</sub>H<sub>22</sub>O<sub>11</sub>), que é indiscernível do açúcar ao olhar humano. Ou seja, se Pascal se tivesse encaminhado ao laboratório de Química munido da sacarina com o intuito de conduzir uma série de experiências, pensando que estaria a usar sucrose, estas teriam falhado devido às diferentes propriedades da sucrose, e teria de concluir que a satisfação do seu desejo por açúcar não foi garantida pela verdade de *p*.

Cremos que a falência do princípio de Ramsey na versão de Whyte se deixa explicar pelo facto de Whyte não ter considerado o carácter *desacoplado de uma representação* relativamente ao estado imediato do sistema (conforme exigido na aplicação da teoria interactivista a representações mentais [cf. §II.2.]). Mas não seria necessário que Whyte tivesse conhecimento desta teoria para chegar à mesma conclusão. De facto, a solução para este problema poderá ser encontrada no trabalho do pragmatista inglês Wright, que terá argumentado num sentido próximo muitos anos antes:

«Não servirá como resposta apontar os objectivos particulares de diferentes indivíduos, pois o sucesso na sua realização não nos oferece um padrão universal de verdade, mas contaria como verdadeira qualquer ideia que assistisse qualquer indivíduo na satisfação dos seus desejos.»<sup>93</sup> [Wright 1913:606]

Porém, desconhecendo o trabalho de Wright, passaram-se anos até que Jérôme Dokic & Pascal Engel [2002 e 2005], atendendo a exemplos como o anterior (de facto, o exemplo de Pascal é em parte uma paráfrase do deles [2002:48]), vieram fazer jus à ideia de Wright, tendo reformulado o princípio de Ramsey do seguinte modo:

---

<sup>92</sup> «A belief's truth condition is that which guarantees the fulfilment of any desire by the action which that belief and desire would combine to cause.»

<sup>93</sup> «It will not answer to cite the particular aims of different individuals, as success in their attainment affords no universal standard of truth, but would mark as true *any* idea that assisted *any* individual in fulfilling his desire.»

«As condições de verdade de uma crença são aquelas que garantem o sucesso de uma acção baseada naquela crença, sejam quais forem os desejos motivadores subjacentes.»

Esta nova versão<sup>95</sup> permite-nos dizer que crenças com conteúdos verdadeiros seriam aquelas que não seriam susceptíveis a revisões ditadas pela falência de acções guiadas por elas, pois essas acções nunca seriam desse modo frustradas – contrariamente ao que aconteceria com a possível acção de Pascal se este não tivesse conversado com a química e, por conseguinte, tivesse continuado a acreditar em *p*’ “O conteúdo da taça é açúcar”, tentando eventualmente realizar (e.g.) certas experiências culinárias como caramelizar a sacarina, e vendo as suas experiências frustradas.

Concluimos então que *crenças falsas também conduzem ao sucesso*, mas nunca sistematicamente. Dokic & Engel chamam a esta tese de *credo pragmatista*: «uma crença falsa só pode ser útil localmente e por acidente. Uma crença falsa não pode guiar apropriadamente o nosso comportamento em todas as situações possíveis» [Dokic & Engel 2002:48-9], enquanto uma crença verdadeira, por oposição, pode.

Ademais, note-se que, em virtude deste *quantificador universal sobre os desejos* (para todos os efeitos equivalente ao desacoplamento estequiométrico), *o princípio torna-se insensível ao contexto da crença*. O contexto é *a fortiori* incluído nas condições de verdade da crença [Dokic & Engel 2005:14].

Noutra ressalva, o princípio diz respeito a crenças *plenas* e não a crenças *parciais*. As condições de verdade da crença (se a crença é verdadeira ou falsa) são independentes do *grau da crença*. Dokic & Engel [2005:9] convidam-nos então a pensar no princípio como aplicando-se a crenças plenas, já que foi também essa a lição inicial, introduzida por Ramsey, que completámos pela análise da tabela-matriz de interacção supra apresentada. Por este motivo, o princípio de Ramsey é pressuposto pela Teoria da Decisão<sup>96</sup>. Com esta adenda podemos colocar de lado críticas como a de Weng Hong Tang [2014:19], que nos diz que o grau da crença parece afectar as propriedades causais da própria crença, tal que uma crença plena é causalmente diferente de uma crença parcial, apesar de terem o mesmo conteúdo. O que a nossa análise demonstra é que as propriedades atitudinais não parecem pesar na determinação das propriedades associadas ao conteúdo propriamente dito, *ainda que possam alterar as propriedades causais do agente* (acerca das quais a própria crença do agente pode versar).

---

<sup>94</sup> «A belief’s truth conditions are those that guarantee the success of an action based on that belief whatever the underlying motivating desires»

<sup>95</sup> À qual nos iremos referir doravante, segue numa direcção diferente da teoria semântica de Ramsey, dado que se refere a condições de sucesso de uma acção e não às causas e efeitos de uma acção, como Ramsey entendia: «The essence of pragmatism I take to be this, that the meaning of a sentence is to be defined by reference to the actions to which asserting it would lead, or, more vaguely still, by its possible causes and effects.» [Ramsey 1927:51]. Bermúdez [2003:§4] é especialmente claro neste ponto.

<sup>96</sup> De facto, parece que o próprio Ramsey estaria disposto a concordar com a análise de Dokic & Engel quando escreveu que «[g]iven a habit of a certain form, we can praise or blame it accordingly as the degree of belief it produces is near or far from the actual proportion in which the habit leads to truth. (...) This is a kind of pragmatism: we judge mental habits by whether they work, i.e. by whether the opinions they lead to are for the most part true, or more often true than those which alternative habits would lead to.» [Ramsey 1929:92, 93-4].

Numa última nota, a estrutura do princípio de Ramsey aplica-se a qualquer forma de representação desacoplada do ambiente. Simon Blackburn [2005:29] enfatiza que a aplicação do princípio não se restringe a veículos de representação *linguísticos* (e o exemplo que dá é de um *mapa cognitivo*), o que está de acordo com o nosso parecer de que mesmo conteúdos primitivos ou representações mínimas podem ser pensados à luz de uma semântica de sucesso entendida *lato sensu*.

### 3. Uma Semântica de Sucesso Interactivo: o Princípio de Ramsey Naturalizado

Whyte não considerou directamente a naturalização da sua versão do princípio de Ramsey [1993], mas autores como Dokic & Engel [2002 & 2005] e Blackburn [2005] consideraram a naturalização das suas. Porém, e infelizmente, só consideraram a naturalização do princípio de Ramsey através da teoria etiológica de funções biológicas, que reprovámos acima [cf. §1.3.]. O motivo pelo qual estes autores se concentraram na teoria etiológica deve-se provavelmente ao facto de os seus proponentes terem-se aproximado, em alguma medida, da própria estrutura do princípio de Ramsey, começando pela própria Millikan. Com efeito, esta disse que:

«O mesmo percepto do mundo pode ser utilizado para guiar qualquer uma de muitas e diversificadas actividades, práticas ou teóricas. O que se mantém o mesmo é que o percepto tem de corresponder a configurações ambientais em acordo com as mesmas regras de correspondência para todas estas actividades.»<sup>97</sup> [Millikan 1989:290, e novamente em 1993:92]

Ou seja, Millikan crê que as condições de verdade de uma determinada representação<sup>98</sup> se mantêm as mesmas independentemente da natureza da acção que se rege por elas<sup>99</sup>. Para Millikan, as representações estão então desacopladas de acções específicas, tal como sucede no caso do princípio de Ramsey. Se a estas aproximações juntarmos ainda o factor popularidade da teoria etiológica, compreende-se que Dokic & Engel tenham considerado que:

«Uma semântica de sucesso é necessariamente uma “teleosemântica”, pois os conteúdos das nossas crenças (e desejos) são determinados, pelo menos em parte, pelas suas funções biológicas, incluindo adaptativas.» [2002:51]

Mas existem ainda outros factores que podem ter alimentado esta asserção e que estão para lá do compromisso da teoria etiológica com a ideia de que as representações mentais (como crenças), «estados internos que acompanham aspectos do nosso mundo, (...) não têm a função de controlar comportamentos particulares»<sup>100</sup> [Sterelny 2003:29]. Conforme já vimos, uma semântica de sucesso tem por base o papel de uma representação na determinação de uma acção, e o mesmo se passa com a teoria teleosemântica da representação [Wrenn

---

<sup>97</sup> «The same percept of the world may be used to guide any of very many and diverse activities, practical or theoretical. What stays the same is that the percept must correspond to environmental configurations in accordance with the same correspondence rules for each of these activities.»

<sup>98</sup> Aqui aceitamos com Millikan e Blackburn que imagens mentais representam, *i.e.* têm conteúdo intencional.

<sup>99</sup> Com a ressalva de que Millikan pressupõe neste excerto a veracidade da Teoria de Verdade por Correspondência [cf. §III.4.].

<sup>100</sup> «internal states that track aspects of our world, but which do not have the function of controlling particular behaviours»

2011:453]. Para além desta semelhança, ambas as teorias concordam ao equacionar as «condições de verdade das crenças com circunstâncias nas quais acções delas resultantes satisfarão desejos»<sup>101</sup> [Papineau 2016:110]. Acrescentando ainda que ambas as teorias parecem consentir com o *slogan* epistemológico de que «a verdade é o alvo da crença», e que «a verdade é aquilo para que as nossas crenças estão direccionadas se as nossas acções são para terem sucesso. Neste sentido, ambas abrem espaço à ideia que a verdade é de algum modo “normativa” para a formação de crenças»<sup>102</sup> [Dokic & Engel 2005:112].

Exposta esta lista de concordâncias, o que se torna evidente e de suma relevância para nós é que tudo isto poderia também dizer-se da tese do tropismo autonómico para a verdade [cf. §II.], de tal forma que deixa de ser necessário entender a semântica de sucesso como “*necessariamente* uma teleosemântica”. Além do mais, temos razões adicionais para não considerar a teoria etiológica como melhor candidata ao lugar de teoria naturalizante da semântica de sucesso. Estas razões prendem-se com o carácter histórico da teoria etiológica previamente criticado [cf. §I.3.]. Recuperando apenas algumas das críticas que são comuns aos proponentes de semânticas de sucesso, é o caso que muitas crenças (e também desejos) surgiram muito tarde na história evolutiva, e, pior, para algumas destas não houve ainda sequer qualquer selecção natural. Relembremo-nos dos exemplos de Bickhard [e.g. 2009] e, mais expressamente, de Whyte [1993:n5] (a possibilidade de um novo mecanismo cognitivo capaz de produzir representações nunca antes instanciadas). Assim, não se pode dizer destas crenças que tenham um propósito evolutivo *à la* teoria etiológica [Blackburn 2005:24]; mas não deixa de ser o caso que estas representações têm condições de verdade. Ademais, o próprio princípio de Ramsey pressupõe, tal como a teoria interactivista, que *as condições nas quais acções têm sucesso são coetâneas com a acção* [Dokic & Engel 2002:54] e não condições de ambientes não-actuais nos quais os nossos mecanismos cognitivos evoluíram, tal como a teoria etiológica pressupõe.

Por estes motivos, Dokic & Engel concluíram que «[nenhuma] função biológica ou considerações evolucionárias podem directamente determinar os conteúdos das nossas crenças e desejos»<sup>103</sup> [*idem*]<sup>104</sup>. E uma vez que Dokic & Engel tinham unicamente em mente a teoria etiológica, no seu segundo ensaio já moderaram o seu optimismo face à naturalização etiológica, apresentando o princípio de Ramsey como *independente* de uma teoria naturalista do conteúdo [2005:12]. Porém, é possível argumentar a favor de uma naturalização interactivista, sabendo que a teoria interactivista não é afectada pelas críticas ao historicismo da teoria etiológica.

## 4. A Natureza Epistémica e Ontológica da Verdade

Outro motivo para considerar positivamente a naturalização interactivista prende-se com a estrutura explicativa do sucesso das acções e com o papel que a verdade desempenha nestas

---

<sup>101</sup> «[Success Semantics] agrees with teleosemantics in equating the truth conditions of beliefs with circumstances in which resulting actions will satisfy desires.»

<sup>102</sup> «Both success semantics and teleosemantics give sense to the familiar claim that “truth is the aim of belief”: truth is what our beliefs are directed to if our actions are to succeed. In this sense, both make room for the idea that truth is in some sense “normative” for belief formation.»

<sup>103</sup> «[no] biological functions and evolutionary considerations can directly determine the contents of our beliefs and desires»

<sup>104</sup> Dokic & Engel [2002:52] e Blackburn [2005:24] pensaram noutras formas de naturalizar semânticas de sucesso que não envolvem uma *determinação directa dos conteúdos pelas suas funções biológicas* (sem por isso abandonarem o contexto da teoria etiológica), mas aqui iremos preocupar-nos unicamente com esta possibilidade, que cremos ser compatível com a teoria interactivista.

explicações. A estrutura da explicação do sucesso das acções/interacções toma contornos semelhantes em semânticas de sucesso e na tese do tropismo autonómico para a verdade. Seguindo a estrutura delineada por Papineau para as semânticas de sucesso [1990:26]:

1. X deseja que  $p$ .
  2. X crê que uma certa acção permitirá obter  $p$ .
  3. Esta crença é verdadeira.
- $\therefore$  X obtém  $p$ .<sup>105</sup>

Esta estrutura dá ainda conta da importância da falência das nossas acções para a revisão das nossas crenças. Com efeito, podemos aprender que uma dada representação é falsa precisamente quando as nossas acções guiadas por essa representação falham<sup>106</sup>. A possibilidade de *detecção de erro* por parte do sistema (registada quando a utilidade esperada não é realizada) é fundamental para o Interactivismo [Bickhard 2009, Miłkowski 2015:7-8]. Esteja consciente dos seus erros representativos ou não, é não obstante possível abordar do mesmo modo a inabilidade de um organismo lidar com um determinado ambiente por respeito a determinada característica (e.g. uma característica ambiental relativamente à qual o organismo não se adaptou e devido à qual o organismo apresentará, enquanto tal, e muito provavelmente, inaptações), o que nos permite corroborar esta estrutura também no contexto do modelo mínimo de representação.

Todavia, esta estrutura poderá receber diferentes interpretações consoante a Teoria da Verdade subscrita pelo seu proponente. Tanto Ramsey quanto Whyte estão comprometidos com a tese de que «a verdade é a propriedade de uma crença que é suficiente para que adquiramos o que queremos quando agimos a partir dela»<sup>107</sup> [Whyte 1990:149] e que nada mais há a dizer acerca da natureza da verdade: a verdade é *redundante*. Neste sentido, Peirce parece ter antecedido a teoria de Ramsey e Whyte na medida em que para ele «a verdade não é nada mais nem menos que o carácter de uma proposição que consiste nisto, que crer na proposição iria, com suficiente experiência e reflexão, levar-nos a uma conduta que tenderia a satisfazer os desejos que então teríamos»<sup>108</sup> [Peirce 1903:CP:5.375, n2]<sup>109</sup>.

Bermúdez [2003] e Godfrey-Smith [1994], por sua vez, concordam que uma semântica de sucesso poderá ser independente de uma Teoria da Verdade, seja qual ela for. Porém, uma semântica de sucesso pode ser compatível, senão mesmo indissociável, de uma teoria de verdade por *correspondência*. Isto torna-se claro a partir da citação de Millikan, tendo sido expressamente aferido por Godfrey-Smith:

«Presumivelmente, quando agimos sob uma crença, nós teremos sucesso ou falharemos de acordo com o modo como as coisas são no mundo. Então também poderíamos dizer: a crença é verdadeira se certas condições no mundo são obtidas, condições nas quais acções guiadas por essa

---

<sup>105</sup> «1. X desires that  $p$ . / 2. X believes that a certain action will bring about  $p$ . / 3. This belief is true. /  $\therefore$  X brings about  $p$ .»

<sup>106</sup> Note-se que nada se diz relativamente à possibilidade de crenças falsas poderem conduzir acções de sucesso. Tudo quanto se pode concluir relativamente a esta possibilidade é que o sucesso de acções conduzidas por crenças falsas não constitui motivos para revisão das crenças.

<sup>107</sup> «truth is the property of a belief that suffices for you getting what you want when you act on it».

<sup>108</sup> «For truth is neither more nor less than the character of a proposition which consists in this, that belief in the proposition would, with sufficient experience and reflection, lead us to such conduct as would tend to satisfy the desires we should then have.»

<sup>109</sup> A *Teoria da Verdade como Redundância* não será explorada neste ensaio.

crença têm sucesso. Parece que temos aqui o material para uma teoria de verdade por correspondência (...)» [1994:260]

Não será então por acaso que Dokic & Engel [2002 e 2005] abandonam a *teoria de verdade enquanto redundância* de Ramsey e Whyte para se aproximarem de uma teoria correspondentista ao dizerem que as condições de verdade não são «nada para além do estado de coisas correspondendo à crença»<sup>110</sup> [2002:47]. Contudo, a adopção de tal teoria permite a aproximação de uma semântica de sucesso à tese do tropismo autonómico para a verdade, pois abre espaço ao enquadramento nomológico (e.g. causal) dos conteúdos por respeito aos seus veículos de representação (em virtude dos quais se dá a correspondência)<sup>111</sup>. Afinal, um tropismo autonómico para a verdade é uma correspondência entre as interacções antecipatórias (ou representações) dos organismos e certos estados de coisas no Mundo, sendo que ambos os *relata* têm propriedades nomológicas definidas, e de tal forma que o «sucesso em providenciar orientação depende não-acidentalmente duma relação entre esta estrutura [representação] ela mesma e aquilo com respeito ao qual a acção é guiada (o que é representado)»<sup>112</sup> [Gładziejewski 2016:24].

Repare-se ainda que uma Teoria da Verdade compatível com uma semântica de sucesso confere *estabilidade* ou *supra-assertibilidade* (*superassertibility*) aos conteúdos verdadeiros, entendendo a estabilidade da verdade como a qualidade de uma proposição verdadeira que faz com que ela seja asserível como verdadeira em qualquer tempo (por vezes necessitando a transformação apropriada do tempo e modo da proposição) [cf. Edwards 2016:2]. E isto é possível devido ao quantificador universal sobre os desejos, que nos permite uma identificação robusta dos conteúdos das representações, querendo com isto dizer que os conteúdos são identificados com «condições invariantes no mundo que garantem o sucesso seja qual for o objectivo perseguido»<sup>113</sup> [Dokic & Engel 2002:47], e devendo-se em última instância esta invariância às propriedades, também elas invariantes, do veículo de representação ou interacção antecipatória.

Outra característica de uma Teoria de Verdade compatível com a tese do tropismo autonómico para a verdade é que deverá ter elementos pluralistas, entendendo-se por uma *teoria pluralista acerca de verdade* uma teoria com elementos de várias teorias usualmente apresentadas como distintas e alternativas (e.g. no caso específico, elementos correspondentistas e *coerentistas*, devendo-se os elementos coerentistas ao facto de que as interacções antecipatórias e os seus conteúdos têm as suas propriedades definidas *relacionalmente*, e que a verdade, sendo uma propriedade dos conteúdos, não será excepção). O sucesso das interacções deve-se, como já vimos previamente, tanto à relação do sistema com o ambiente, como às relações entre as partes que constituem o interactor (no caso das propriedades de uma crença, entre as quais o seu valor de verdade, as propriedades doutra(s) crença(s), incluindo os valores de verdade des-toutas).

Além disto, porque há relações causais e ecológicas de carácter extrínseco entre os interactores e os seus *relata* no mundo, dizemos que o valor de verdade de uma proposição é também uma propriedade extrínseca: alterações no meio ambiente podem causar alterações no valor de

---

<sup>110</sup> «are nothing but the state of affairs corresponding to the belief»

<sup>111</sup> Interessantemente, também para Ramsey [cf. 1927:163-4] são as propriedades causais das crenças verdadeiras e falsas e mais precisamente as suas causas e efeitos que são de relevo [cf. Methven 2015:153], contudo, Ramsey não adoptou uma teoria de verdade por correspondência.

<sup>112</sup> «a given action was representation-guided if it was guided by a structure (representational vehicle) whose success in providing action guidance non-accidentally depends on a relation between this structure itself and that with respect to which the action is guided (what is represented).»

<sup>113</sup> «invariant conditions in the world that guarantee success whatever goal is pursued.»

verdade de uma representação sem que as restantes propriedades do veículo de representação sofram alterações. Assim, como as propriedades dos veículos representacionais e a sua semântica são condicionadas pelo holismo da fitness/autonomia e pelo holismo semântico, respectivamente, dizemos então que a verdade de uma proposição é uma propriedade *relacional*, entendendo propriedades relacionais como «propriedades que cada entidade tem ou adquire devido às suas relações extrínsecas com outras entidades no seu ambiente»<sup>114</sup> [Santos 2015b:25; cf. ainda Edwards 2016].

Para finalizar este capítulo perguntamo-nos, mas deixamos em aberto, como é que o princípio de Ramsey é afectado quando os conteúdos de uma representação são apenas *parcialmente* ou *aproximadamente verdadeiros*, e não *stricto sensu* verdadeiros ou falsos. Como a detecção de falsidade se dá pela falência de uma acção conduzida por uma representação, acções que não falham não foram falsificadas e poderão ser, enquanto tal, verdadeiras, apesar de nada nos permitir aferir com certeza que são verdadeiras. O seu sucesso enquanto *stricto sensu* falsas, porém, poderá ser explicado se *parte* dos conteúdos proposicionais forem verdadeiros. A título de exemplo: dizem a Pascal que, *f* “O sumo está no frigorífico, na cozinha”, e Pascal, que não tem motivos para desconfiar do testemunho, age de acordo com *f* para satisfazer o seu desejo por sumo, dirigindo-se ao frigorífico da cozinha. Porém, o sumo não estava no frigorífico, mas nas mãos de Newton, que se servia do sumo no metro quadrado em frente ao frigorífico, e Pascal consegue ainda assim satisfazer o seu desejo, partilhando o sumo com Newton. Uma noção apropriada de aproximação à verdade ficará por desenvolver [cf. Niiniluoto 1999].

## 5. A Natureza dos Conteúdos de Volições e a Subordinação da sua Funcionalidade à Autonomia

Autores como Papineau defendem ainda que as Semânticas de Sucesso dependem de uma teoria dos conteúdos dos desejos, pois uma Semântica de Sucesso «explica a verdade para as crenças, apenas assumindo uma noção de satisfação para os desejos. Contudo, satisfação é uma noção representacional tanto quanto verdade, e portanto deve ela mesma ser explicada por uma teoria filosófica da representação adequada»<sup>115</sup> [Papineau 1993:70-1]. Aqui não dedicaremos muito tempo à apresentação de uma teoria dos conteúdos dos desejos, mas avançaremos ainda assim com alguns comentários que nos poderão ajudar a pensar possíveis contra-exemplos às teses deste ensaio.

As teorias de conteúdos de desejos são usualmente pensadas à parte de representações como crenças, pois é costumeiro pensar que estados mentais cognitivos, como as crenças, diferem de estados volitivos ou motivacionais (de que são exemplo desejos), porquanto os primeiros estão dirigidos à verdade e os últimos não [cf. Engel 2004:77]. Outra forma de pensar esta direcionalidade deve-se a G.E.M. Anscombe, que fez uso de uma metáfora esclarecedora. Diz-nos ela que crenças e desejos diferem na sua “directção de encaixe” (*direction of fit*), na medida em que as crenças se adaptam ao mundo, enquanto é o mundo que se adapta aos desejos. Robert Stalnaker pô-lo mais explicitamente:

---

<sup>114</sup> «relational properties are properties that each entity has and acquires due to its extrinsic relations with other entities in its environment»

<sup>115</sup> «[SS] explains truth for beliefs, only by assuming the notion of satisfaction, for desires. Yet satisfaction is as much a representational notion as truth, and so ought itself to be explained by an adequate philosophical theory of representation.»

«Crença e desejo (...) são estados disposicionais relativos de um agente potencialmente racional. Desejar que P é estar disposto a agir de maneiras que tendem a provocar P num mundo em que as crenças do agente, quaisquer que elas sejam, sejam verdadeiras. Acreditar que P é estar disposto a agir de maneiras que tendem a satisfazer os nossos desejos, quaisquer que eles sejam, num mundo em que P (juntamente com as restantes crenças do agente) fossem verdadeiras.» [Stalnaker 1987:15]

De acordo com este modo de pensar, *os desejos não têm condições de sucesso como as crenças*, mas, sim, *condições de satisfação* [e.g. Engel 2004:79], pelo que é habitual concluir que os desejos não podem ser verdadeiros ou falsos. Esta caracterização opõe-se à da teoria interactivista, pois a teoria interactivista está comprometida com a tese de que todas as interações funcionais têm conteúdos, definidos como condições de verdade, identificados como condições de sucesso das interações. Ou seja, o Interactivismo dirá que desejos funcionais são verdadeiros.

De facto, de acordo com o Interactivismo, podemos *definir os conteúdos dos desejos pelas suas condições de satisfação* e manter, ainda assim, que *os desejos verdadeiros são aqueles cujas condições de satisfação, de acordo com o tropismo autonómico para a verdade, mantêm a autonomia*. Neste sentido, também um desejo pode ser falsificado se, p.ex., a volição do biosistema for direccionada a um objecto que não contribui para a preservação da sua autonomia. Com esta adenda, podemos então concordar com a ortodoxia no sentido em que aceitamos que *a satisfação de desejos não faz deles verdadeiros ou falsos*, acolhendo ainda assim os conteúdos dos desejos no seio da teoria interactivista de acordo com a qual também podem ser verdadeiros ou falsos.

Uma teoria interactivista dos conteúdos dos desejos, de acordo com a qual estes são verdadeiros ou falsos consoante a sua funcionalidade subordinada à preservação da autonomia dos seus sujeitos, distingue-se imediatamente das semânticas de sucesso existentes e doutras teorias acerca dos conteúdos dos desejos, pois tem por consequência que *as condições de verdade de uma determinada representação se tornam independentes da viabilidade dos desejos*, o que se torna evidente na eventualidade dos desejos não promoverem a autonomia do sistema (e.g. desejos suicidas ou a *resposta normal* à presença de *estímulos supranormais* que conduz alguns *H. sapiens* à obesidade mórbida [cf. Barrett 2010]).

Os desejos são, como previamente apontado, representantes de normas autonómicas, não coincidindo necessariamente com estas<sup>116</sup>. É de facto possível, e por vezes recomendável, que não satisfaçamos os nossos desejos (esta é uma das razões pela qual Papineau [e.g. 1990 & 2000] sugere que as semânticas de sucesso necessitam de uma teoria naturalista da função dos desejos). Retomando o último exemplo, o desejo de evitar a obesidade (tendo em vista evitar problemas consequentes) pode ser tomado como uma volição de segunda ordem, *sensu* Harry Frankfurt, que é funcional na medida em que preserva a autonomia por oposição ao desejo de continuar a comer na presença de estímulos supranormais<sup>117</sup>. A subordinação da função dos

---

<sup>116</sup> Do ponto de vista biológico é ainda interessante observar que parece ser o caso que a funcionalidade de estados volitivos mais primitivos como (e.g.) a sede, por oposição à funcionalidade de estados volitivos mais complexos como (e.g.) o desejo de experienciar nostalgia ao ouvir a interpretação de Abbado & Kožena da *Ich bin der Welt abhanden gekommen* no silêncio de uma noite de Verão com vista para os Dolomitas ao sabor de um bom Scotch e em perfeita companhia, está mais dependente de um acoplamento estequiométrico adequado do que de um desacoplamento. Aqui não desenvolveremos mais esta problemática.

<sup>117</sup> Conforme nos foi apontado por Margarida Hermida, o caso de Pascal e da taça de sacarina também tem relevância aqui, já que a sacarina foi introduzida no mercado de modo a satisfazer o desejo por sabores adocicados e de preservar a autonomia dos agentes, enquanto o nosso desejo por coisas doces no



desejos à autonomia permite então explicar porque é que *crenças verdadeiras podem não obstante conduzir à morte dos agentes*. Que o possam fazer é apenas o resultado de acções conduzidas de modo a satisfazer desejos falsos.

## 6. Contraexemplos à Semântica de Sucesso Interactivista e Respostas

Tendo por base os avanços anteriores podemos agora enfrentar com maior facilidade alguns contra-exemplos à tese do tropismo autonómico para a verdade e às semânticas de sucesso.

### (a) *Subdeterminação na atribuição de conteúdos.*

Apesar das semânticas de sucesso não serem teorias acerca da atribuição de conteúdos, é costume apontar-lhes a crítica de que o conteúdo de uma representação é subdeterminado pelo comportamento, devido ao facto de que o mesmo comportamento se pode seguir de vários conteúdos e de que vários comportamentos se podem seguir de uma mesma representação. A proposta das semânticas de sucesso para resolver o problema da subdeterminação de um conteúdo pelo comportamento passa por especificar o conteúdo através da determinação de todas as condições relevantes para o sucesso de todos os comportamentos possíveis que concordam com uma dada representação. Nas palavras de Bermúdez, «o constituinte “correcto” do conteúdo é aquele que melhor se projecta numa gama apropriada de padrões de comportamento actuais e contrafactuais»<sup>118</sup> [Bermúdez 2003:108]. Assim, o problema da determinação de conteúdos pode ser trabalhado dum ponto de vista contrafactual [Dokic & Engel 2002:49], sendo possível isolar as condições de verdade de uma representação «considerando muitas variações relevantes ou transformações de uma dada situação na qual a crença desempenha um papel comportamental. As condições de verdade da crença irão eventualmente emergir como a invariante de todas essas transformações»<sup>119</sup> [Dokic & Engel 2002:49].

Contudo, há que reconhecer a seguinte limitação epistémica ao método contrafactual: há um conjunto potencialmente infinito de acções e circunstâncias em que a mesma representação pode interagir. Apesar de tudo, esta limitação não derrota a semântica de sucesso. E tanto quanto o seu apelo contra-argumentativo depende desta limitação epistémica, deveremos então apontar que agentes humanos parecem ser especialmente eficazes na identificação das circunstâncias de relevo na determinação do sucesso das suas acções. Porque esta problemática é para todos os efeitos a mesma que a da determinação do ambiente de referência da fitness, cremos ser útil trazer à consideração o caso da determinação da fitness de um urso polar sem pelo [o exemplo é de Goode & Griffiths 1995]. Compreendemos fácil e intuitivamente que é devido à falta de pelo que o urso polar não sobreviverá ao inverno do Ártico, apesar de existirem outras características e condições das quais os efeitos negativos da falta de pelo também dependem, em acordo com a lógica do holismo da fitness (e.g. o ritmo de batimento cardíaco, a

---

nosso passado ambiente de selecção natural não teria por objecto a sacarina, mas outros açúcares que ocorrem naturalmente em alimentos ricos em hidratos de carbono e altamente calóricos.

<sup>118</sup> «The “correct” content-constituent is the one that best projects across an appropriate range of actual and counterfactual patterns of behavior.»

<sup>119</sup> «by considering many relevant variations or transformations of a given situation in which the belief plays a behavioural role. The belief’s truth conditions will eventually emerge as the invariant of all these transformations.»

pressão sanguínea, a alimentação do urso, outras condições ambientais como a humidade e o vento, etc).

(b) *Sobrecarga cognitiva.*

A crítica de sobrecarga cognitiva (*cognitive overload; overburdening of cognition*) foi colocada independentemente por John Perry [1993] e Robert Brandom [1994], e deve-se à formulação do princípio de Ramsey segundo a qual conteúdos verdadeiros são aqueles que *garantem* o sucesso das nossas acções.

Diz-nos Perry que seria irrealista «pressupor que os conteúdos das nossas crenças fixam todas as circunstâncias relevantes para o sucesso da nossa acção»<sup>120</sup> [1993:202, conforme citado por Dokic & Engel 2005:14]. A crítica é simples: parece que muitas das circunstâncias que garantem o sucesso das nossas acções são independentes do nosso estado cognitivo-epistémico. Contudo, o princípio de Ramsey parece dizer-nos que «[s]e todas as possíveis falhas forem explicadas por crenças falsas, a crença verdadeira correspondente deve estar presente quando temos sucesso. Assim, quando eu alcanço o copo, devo acreditar que as forças da gravidade são tal como têm de ser para as coisas decorrerem bem» [*idem*], quando certamente não é o caso que nós a cada momento mantemos (e.g.) a crença verdadeira de que a aceleração gravítica é  $\approx 9,8 \text{ m/s}^2$ , nem outras crenças acerca de todas as condições normais subjacentes às nossas acções<sup>121</sup>.

Brandom acredita que o facto de muitas crenças diferentes estarem envolvidas no sucesso de uma só acção – um problema que Blackburn [2005:23] identifica como o já mencionado *holismo do mental* – não representa um problema de maior, bastando que o princípio de Ramsey fosse estendido ao conjunto de todas as crenças responsáveis pela acção [Brandom 1994:175]. Contudo, o derradeiro problema advém da observação de que o sucesso de uma acção requereria do agente que este também mantivesse crenças acerca do *não impedimento* da sua acção por esta ou por aquela circunstância [Brandom 1994:177]. A gravidade desta observação deriva do facto de que o conjunto das crenças acerca da inexistência de impedimentos à acção tem, potencialmente, infinitos membros. Pascal poderia não acreditar activamente em *p*: “Não existe um assassino profissional desejoso de me matar escondido atrás do balcão onde presentemente repousa a taça de açúcar do qual me vou servir para adoçar o meu café” quando teve sucesso ao adoçar o seu café, mas não obstante parece óbvio que o facto de não existir tal assassino escondido atrás do balcão é uma das condições que explica o sucesso de Pascal. Quantas outras crenças acerca da ausência de impedimentos *sui generis* poderiam formular-se?

Concluem Perry, Brandom, e Blackburn<sup>122</sup>, que o princípio de Ramsey requer demasiado das nossas capacidades cognitivas. Pura e simplesmente, nós não mantemos activamente tantas

---

<sup>120</sup> «[L]et us first note how unrealistic it would be to suppose that the content of beliefs fix all the circumstances relevant to the success of our action.»

<sup>121</sup> Godfrey-Smith coloca de lado este tipo de exemplo observando que «[u]niversally necessary parts of success-conditions can be factored out of truth-conditions just because they are universal.» [Godfrey-Smith 1994:262]. Nós acreditamos que estas condições universais de que ele fala podem constituir somente o ambiente habitual de selecção das nossas disposições, mas que estão longe de serem de facto universais. Ao fim de um ano a viver na Estação Espacial Internacional, podemos dar por certo que o astronauta Scott Kelly e o cosmonauta Mikhail Kornienko não mantêm, ao longo do seu dia, e para quase todas as suas acções, a crença de que  $g \approx 9,8 \text{ m/s}^2$ .

<sup>122</sup> Blackburn [2005:23-4] aceita a crítica de Brandom, em virtude da qual modera a sua versão de uma semântica de sucesso.

crenças aquando da concretização das nossas acções que têm sucesso, pelo que o princípio não pode estar certo.

Dokic & Engel [2005] sugeriram uma solução para este problema. Começam por sugerir que pensemos a distinção entre crenças implícitas e explícitas. Quando explicamos o sucesso das nossas acções, dizem, fazêmo-lo remetendo-o para crenças *explícitas*, aquelas acerca das quais o princípio de Ramsey deve primeiramente versar. Blackburn [2005:30] traça esta mesma distinção quando diz que não incluímos nas *explicações* do sucesso das nossas acções crenças implícitas acerca de condições gerais de sucesso (novamente, tome-se o exemplo da aceleração gravítica), devido à natureza *diferencial* ou *contrastante* da explicação [cf. ainda Godfrey-Smith 1994:262].

Dokic & Engel [2005:19] acrescentam ainda que muitas vezes agimos no desconhecimento da existência de impedimentos, e nessas situações a própria *acção* torna-se uma *fonte de conhecimento* acerca das suas próprias condições de sucesso (o seu sucesso tem como consequência lógica um ganho de conhecimento de que tais impedimentos não existiam). Agimos *como se* não existissem tais impedimentos, e este “como se” afigura-se como crenças implícitas «no sentido de que se [as crenças] fossem formadas, elas seriam directa ou indirectamente justificadas pela experiência de acção do agente»<sup>123</sup> [Dokic & Engel 2005:19]. Dokic & Engel chamam a isto o “*princípio do fechamento pragmático*” (*principle of pragmatic closure*), segundo o qual «se estou intencionalmente a fazer *p*, e *q* implica que não estou, então eu sei pelo menos implicitamente que *q* não é o caso»<sup>124</sup> [*idem*].

No nosso entender, o princípio do fechamento pragmático é ainda epistemicamente demasiado exigente ou optimista. O princípio diz-nos que, se temos sucesso a fazer *p*, sabemos implicitamente que as condições que nos teriam impedido de fazer *p* não ocorreram, mas o princípio deveria dizer algo acerca da natureza dos impedimentos<sup>125</sup>. Todavia, a situação descrita não requer que possamos vir a *conhecer* as condições de sucesso *específicas* das nossas acções (incluindo aquelas que impediriam a sua concretização, como *q*), nem mesmo implicitamente. Do mesmo modo, podemos estar intencionalmente a fazer *p*, não saber que *q* implica e justifica o insucesso da nossa acção de fazer *p*, falhar na acção, e não ter capacidade de *conhecer q*, *nem mesmo implicitamente*. Tudo quanto poderemos dizer dessa situação é que houve impedimentos, não *quais*.

Creemos que a exigência epistémica que leva à crítica da sobrecarga cognitiva pode ser dispensada com o enquadramento da estrutura do princípio de Ramsey no contexto da tese do tropismo autónómico para a verdade, pois, de acordo com esta, a natureza dos conteúdos é sempre, primeiramente, implícita e subdóxástica (considere-se qualquer antecipação organísmica não-cognitiva como exemplo), não necessitando estes jamais de serem explicitamente formados. Afinal, pode ser dispendioso formar crenças explícitas e, recordando o *princípio de economia da natureza* expresso no *slogan* de Dennett [e.g. 2017:336], há imenso sucesso na natureza (e, por sinal, sucesso bastante sofisticado) que não depende de crenças explícitas.

---

<sup>123</sup> «Most of the relevant beliefs are implicit, in the sense that *if* they were to be formed, they would be directly or indirectly justified by the agent’s *experience of acting*.»

<sup>124</sup> « If I am intentionally doing *p*, and *q* implies that I am not, I at least implicitly know that *q* is not the case.»

<sup>125</sup> Não temos a certeza de que Dokic & Engel concordariam com esta leitura do princípio do fechamento pragmático, de acordo com a qual o princípio parece sugerir que tudo é conhecível pelo menos implicitamente. Cf. Vlerick 2017 para uma defesa de que tudo é em princípio epistemicamente acessível.

Ademais, ao radicar as condições de sucesso nas propriedades das interações antecipatórias, podemos tratá-las de acordo com uma análise causal, o que nos permite identificar as propriedades de uma representação que são *mais relevantes* para a determinação das condições de sucesso e de impedimento, *i.e.* aquelas que, por fazerem a diferença, integram as explicações. Notoriamente, esta é a mesma estratégia com que adquirimos uma compreensão intuitiva do insucesso de um urso polar sem pelos [cf. §III.6.(a)]. É devido à queda total de pelo (causada por uma doença) que o urso polar não sobreviverá ao inverno do Ártico, mesmo que a falta de pelo isoladamente não explique o insucesso do urso polar. Porém, ninguém dirá que a fitness do urso, com ou sem pelos, está “sobrecarregada” de pressupostos. Ela é simplesmente determinada por esses pressupostos, devido à sua natureza inerentemente relacional. *Mutatis mutandis*, a natureza da dependência holística do mental também não implica a sobrecarga da cognição, pois o sucesso de todos os nossos estados mentais representantes e explícitos enquanto guias de acção deve-se parcialmente a conteúdos implícitos, extrinsecamente determinados. Também Dokic & Engel (e Robert Nozick antes deles<sup>126</sup>) chegam a uma conclusão próxima desta ao verificarem que:

«[A] objecção da sobrecarga cognitiva é respondida distinguindo entre conhecimento implícito e explícito. (...) [A] objecção (...) não vinga precisamente porque os agentes estão normalmente adaptados ao seu ambiente. A adaptação não é puramente uma relação externa entre um agente e o seu ambiente, como se o primeiro acontecesse estar “adaptado” ao segundo. (...) O sucesso das nossas acções reflecte-se no estado cognitivo do sujeito, mesmo que apenas implicitamente, porque as contribuições do agente e da Mãe Natureza estão tão interligadas que é impossível distingui-las.»<sup>127</sup> [Dokic & Engel 2005:20].

(c) *O sapo de Miłkowski*<sup>128</sup>.

O sapo de Miłkowski forma uma representação quando uma mosca atravessa o seu campo de visão e age de modo a capturar a mesma para sua nutrição. No acto de projecção da sua língua preênsil, um passeriforme que até então estava fora do seu campo de visão, com uma maior destreza e agilidade, rouba a presa do sapo, causando a falha da acção do sapo. O que dizer então da relação entre a representação aparentemente verdadeira do sapo e a sua falha na acção? Cremos ser possível dizer três coisas: (1) contamos a não-inclusão do passeriforme na representação do sapo como um erro representativo (*i.e.* como falsa), (2) contamos o não-representado como não pesando no valor de verdade do representado, ou (3) dizemos do sapo

<sup>126</sup> «Evolution may have instilled in us phylogenetically information about, and patterns of behavior suitable to, stable facts of our evolutionary past. We do not have to think explicitly about (or even know about) the regular alternation of day and night; our bodily rhythms do that job for us, as jet-lagged travelers discover. Some facts are stable enough so that each organism need not arrive at a knowledge of them itself; evolution will have wired that in. Rationality may have the evolutionary function of enabling organisms to better cope with new and changing current situations or future ones that are presaged in some, possibly complex, current indications. That rationality can do this is itself one of the stable facts, not a fact we need explicitly to know. All we need is to have it built into us, as the presence of gravity is. Evolution employs and builds mechanisms around constant and stable environmental features. It is not just that gravity is a constraint upon some characteristics – for instance, size – but that gravitational force is utilized in the working of some processes.» [1993:120].

<sup>127</sup> «[T]he objection of cognitive overload is answered by distinguishing between implicit and explicit knowledge. (...) [T]he objection of cognitive overload does not stand precisely because agents are normally adapted to their environment. Adaptation is not a purely external relation between an agent and its environment, as if the former happened to “fit” the latter. (...) Our actions’ success conditions reflect themselves in the subject’s cognitive state, if only implicitly, because the agent’s contribution and that of Mother Nature are so intertwined that it is impossible to tell them apart.»

<sup>128</sup> Devo este exemplo a Marcin Miłkowski.

que tem uma representação *incompleta* do mundo relevante para as suas acções. Se (1), o princípio de Ramsey está salvo. Mas o que sucede em (2) e (3) não nos parece tão óbvio.

A questão em causa deve-se, por um lado, à natureza da relação entre entidades reais mas não-representadas e os agentes, e, por outro, aos limites dos poderes causais de representações incompletas e à possibilidade de considerarmos representações incompletas como *stricto sensu* falsas. Neste sentido, a resposta a este contra-exemplo pode tomar contornos similares aos da estratégia delineada na objecção anterior relativamente ao princípio do fechamento pragmático. Pessoalmente, parece-nos correcto contar representações incompletas como falsas, sendo que o fazemos independentemente de sabermos se a descrição é incompleta (e.g. porque estamos momentaneamente cognitiva ou epistemicamente cegos, de tal modo que o nosso modelo causal do mundo não quantifica uma intervenção putativa, *sensu* Woodward [2003:45]) ou falsa (e.g. no sentido em que atribui propriedades inexatas ao estado de coisas do mundo). Gostaríamos de argumentar que tudo quanto nos parece relevante para (2) e (3) é saber se a representação descreve ou não correctamente o estado de coisas do Mundo com que o seu portador *interage*, e, evidentemente, o sapo interagiu com o passeriforme, de tal forma que a sua representação não descreve exactamente o estado de coisas do Mundo com que interage. Assim, parece-nos ser o caso que (2) só constituiria uma interpretação razoável se o não-representado não pudesse interagir com o representante, e que (3) também não viola o princípio de Ramsey.

(d) *Representações irrelevantes para a agência.*

Dokic & Engel [2002:50] apresentam uma possível objecção que atribuem a John Campbell. De acordo com Campbell, nós conseguimos representar o espaço independentemente da nossa capacidade de orientação no mundo. Isto é, teremos representações espaciais do mundo que são independentes das nossas acções sobre o mundo, de tal forma que os conteúdos de tais representações não poderão nunca ser dados pelas condições de sucesso de acções guiadas por essas representações<sup>129</sup>. Dokic & Engel não elaboram a crítica de Campbell, pelo que não é claro que tipo de representações espaciais Campbell terá em mente<sup>130</sup>.

Todavia, e esquecendo a crítica particular de Campbell, deveremos levar a sério a existência de conceitos teóricos que, de um modo imediato, parecem irrelevantes para a agência. Este é o caso de representações de espaços não-euclidianos (e.g. como definidos pelas geometrias de Minkowski e de Riemann) que violam a fenomenologia da nossa percepção. Sobre representações deste tipo gostaríamos de dizer duas coisas. Primeiramente, que não é claro para nós que estas geometrias não tenham influência, ainda que indirecta, nas próprias acções dos físicos-matemáticos quando estes procuram (e.g.) demonstrar axiomas de geometria algébrica, ou quando fazemos uso de tecnologias como o GPS que dependem do uso da teoria da relatividade geral, que por sua vez faz amplo uso destas geometrias desafiantes para a nossa percepção. Depois, que qualquer termo teórico pode ser abordado pelo *método de Ramsey-Lewis* [cf. Lewis 1970], de acordo com o qual o significado de termos teóricos é dado indirectamente pelo papel

---

<sup>129</sup> Também Bermúdez [2003:66] crê que não é possível avaliar os conteúdos de criaturas linguísticas em termos de uma semântica de sucesso porque as suas crenças têm pouco contacto directo com acções e desejos. Porém, Bermúdez não substantia o seu argumento apontando termos que não têm ligação a acções e desejos, nem deixa claro porque é que este problema não poderá ser resolvido.

<sup>130</sup> António Zilhão indicou-nos a possibilidade da crítica de Campbell ser de índole Kantiana. Infelizmente não somos peritos em Kant (nem estamos familiarizados com o trabalho de Campbell), pelo que não poderemos fazer jus a esta crítica. Não obstante, gostaríamos de recordar que, como vimos em §II.1., Lorenz sugere uma “naturalização do sistema Kantiano” de acordo com a qual mesmo as “formas puras” e *a priori* da “intuição sensível” são *filogeneticamente a posteriori*, o que nos permitiria trabalhá-las de acordo com a tese do tropismo autónomo para a verdade.

desempenhado pelos termos no conjunto possível de asserções de uma dada teoria, que incluirá frases (também chamadas de *frases de Ramsey*) que poderão ter implicações experimentais variadas, *i.e.*, com efeitos práticos [cf. Maxwell 1962, Hooker & Churchland 1985, e Andreas 2013 para elaborações acerca dos problemas levantados por termos teóricos].

(e) *O anjo-da-guarda de Godfrey-Smith* [1994:262-3].

O anjo-da-guarda de Godfrey-Smith é um familiar do demónio Cartesiano. Ele intervém nas nossas acções de modo a garantir-lhes o sucesso, e poderá fazer isto quando, na sua ausência, as nossas acções não satisfariam os nossos desejos. Porém, um demónio ainda mais demoníaco não necessitaria de intervir somente nessas alturas, podendo alterar também as circunstâncias nas quais, *ceteris paribus*, as nossas acções teriam sucesso<sup>131</sup>. Assim, a existência de anjos-da-guarda teria por implicação que as condições de sucesso das acções não poderiam ser dadas por nenhuma relação específica entre as interacções de agentes e o seu meio, dado que estas seriam impossíveis de determinar. Isto aconteceria porque a existência de tal anjo-da-guarda deixaria-nos incapazes de decidir quais das nossas representações são verdadeiras ou falsas porque incapacitaria o uso de uma análise contrafactual que se estenderia à incapacitação da nossa aprendizagem por tentativa e erro. O anjo-da-guarda constitui-se deste modo como um cenário de cepticismo Cartesiano contra o princípio de Ramsey. Não responderemos a esta crítica aqui<sup>132</sup>.

(f) *A galinha de Russell*<sup>133</sup>.

A galinha de Russell constitui um exemplo de cepticismo Humeano. Russell descreve a seguinte situação:

«O homem que alimentou a galinha todos os dias ao longo da sua vida acaba por torcer-lhe o pescoço, demonstrando que visões mais refinadas acerca da uniformidade da natureza teriam sido úteis à galinha.»<sup>134</sup> [1959:63]

Este exemplo demonstra que a sorte pode demorar muito tempo, mas não para sempre. É o próprio Russell quem o sugere, ao aludir ao carácter não uniforme da natureza e à falta de adequação da crença da galinha a esse respeito<sup>135</sup>. Assim, este exemplo não rompe com o credo pragmatista de que só representações verdadeiras permitem agir *sistematicamente* com sucesso. A relação interna entre a veracidade de uma representação e o seu sucesso interactivo permanece indisputada.

Contra nós, um céptico poderá acrescentar que sorte eterna é ainda assim possível. Nós respondemos que tal sorte é compatível com a tese de que a probabilidade de um biosistema com uma representação falsa sobreviver em diferentes ambientes agindo de acordo com a

---

<sup>131</sup> Um demónio maligno poderia ainda, contrariamente ao anjo-da-guarda, converter condições que, *ceteris paribus*, garantiriam o sucesso das nossas acções em condições desfavoráveis à satisfação dos nossos desejos.

<sup>132</sup> Confessamos o nosso desinteresse epistémico por cenários deste género, que incluem outros contra-exemplos como “encéfalos em cubas” (*brain in a vat*), simulacros como o “Matrix”, etc., o que talvez se deva à nossa simpatia com o empirismo e pela “espada laser flamejante do Newton” (*Newton’s flaming laser sword*) [Alder 2004], que nos convida a ignorar aquilo que não pode ser ajuizado pela experiência.

<sup>133</sup> Este exemplo foi-nos primeiramente sugerido por João Cordovil.

<sup>134</sup> «The man who has fed the chicken every day throughout its life at last wrings its neck instead, showing that more refined views as to the uniformity of nature would have been useful to the chicken.»

<sup>135</sup> Um vegetariano poderá discordar, pelo que o exemplo se calhar não é o melhor.

mesma representação tende para 0 (tendência que poderá ser justificada atendendo às propriedades causais das interacções e à probabilidade baixa da sorte eterna). Willard van Orman Quine expõe com humor esta ideia quando nos diz que «[a]s criaturas inveteradamente erradas nas suas induções têm uma tendência patética, mas louvável, para se extinguirem antes de reproduzirem a sua espécie»<sup>136</sup> [Quine 1969:126]<sup>137</sup>. Ford Doolittle diz-nos algo parecido num contexto completamente diferente, mas igualmente relevante para a avaliação da fitness: «quanto mais este planeta portador de vida persistiu como tal, tanto mais provável é que haja uma razão para além da sorte para a sua persistência [enquanto planeta portador de vida]»<sup>138</sup> [Ford Doolittle 2014:420].

(g) *Ilusões positivas.*

Um contra-exemplo comum às semânticas de sucesso é o de que crenças falsas podem dar azo a acções de sucesso. A título de exemplo, este será o caso de crenças acerca de situações de risco e incerteza, nas quais uma atitude optimista parece conferir à acção a confiança necessária para que ela se concretize com sucesso na satisfação de desejos (e.g. um caso em que um organismo sobrestima as suas capacidades de combate, crendo-se invencível [o exemplo é de Dokic & Engel 2002:48, e é mais desenvolvido por Owen Flanagan 2009]). Porém, crenças deste tipo – chamemos-lhes “ilusões positivas” em acordo com a literatura – não são *sistematicamente* boas (afinal, «uma incapacidade de ver as suas próprias fraquezas pode impedir as pessoas de alcançar objectivos importantes e podem colocar a sua vida e segurança em perigo»<sup>139</sup> [Dweck 2009:518]; é fácil imaginar alguém que acredita ser à prova de bala e que por consequência é capaz de avançar em vão contra mísseis balísticos) [cf. Dokic & Engel *op. cit.* e Dweck 2009:518], de tal forma que o credo pragmatista [cf. §III.2.] é preservado contra estes exemplos.

Para além disto, gostaríamos de observar que não é claro que a nossa posição epistémica seja tal que nos será permitido aferir se as crenças em causa têm de facto um valor de verdade negativo. Para efeitos práticos, se as acções de agentes epistémicos guiados por tais “ilusões positivas” tiverem sucesso, então “ilusões positivas” não seriam apenas optimistas e adaptativas, mas verdadeiras, *i.e.* as expectativas ditas “ilusórias” seriam, ao fim e ao cabo, aquelas cujas representações acabariam por se provar verdadeiras sempre que guiassem acções de sucesso. Ou seja, não seriam de todoilusórias [voltaremos a esta questão mais à frente, §IV.6.(a)].

(h) *O suicida de Glock*<sup>140</sup>.

O suicida de Glock acredita que a sociedade fica melhor se se matar, e tem razão para acreditar que se se matar a sociedade ficará melhor, *i.e.* a sua crença é verdadeira. Assim, o caso do suicida de Glock constitui um exemplo em que o suicida tem uma crença verdadeira e satisfaz o seu desejo de morrer. Resta perguntar se o desejo do suicida de Glock é verdadeiro. A resposta imediata poderá ser que não, pois é óbvio que o desejo do suicida não contribui para a sua

<sup>136</sup> «Creatures inveterately wrong in their inductions have a pathetic but praiseworthy tendency to die out before reproducing their kind.»

<sup>137</sup> No contexto da asserção, Quine refere-se ao “espaçamento subjectivo de qualidades”: «Why should our subjective spacing of qualities have a special purchase on nature and a lien on the future? / There is some encouragement in Darwin. If people’s innate spacing of qualities is a gene-linked trait, then the spacing that has made for the most successful inductions will have tended to predominate through natural selection.» [Quine 1969:126]

<sup>138</sup> «[T]he longer this life-bearing planet has persisted as such the more likely there is to be a reason other than luck for its persistence.»

<sup>139</sup> «an inability to see their own weaknesses can prevent people from reaching important goals and can endanger their health and safety»

<sup>140</sup> Este exemplo foi-me sugerido por Hans-Jonnas Glock.

própria autonomia. Porém, o suicídio do suicida de Glock contribui para a autonomia da sociedade. Aliás, é sobre esse pretexto que é, em primeiro lugar, possível dizer que a sua crença de que a sua morte é boa para a sociedade é verdadeira. Na medida em que a perspectiva autonómica abraça a teoria hierárquica dos sistemas [e.g. Christensen & Bickhard 2002:3; cf. §1.2.3.], este exemplo não se lhe contrapõe.

(i) *O monge de Miłkowski* [2015:1]<sup>141</sup>.

O monge de Miłkowski respeita o celibato e enquanto tal não se reproduz. Por consequência, apesar das crenças do monge terem conteúdos potencialmente verdadeiros, haverá pelo menos uma volição de segunda ordem, motivada pelas crenças religiosas do monge, que o leva a não se reproduzir, e que não contribui, enquanto tal, para a preservação da frequência em que os seus alelos surgem na linhagem ao longo do tempo. Novamente, a abordagem hierárquica permitirá perspectivar os aspectos culturais do monge relativamente aos quais poderemos então pesar se o celibato do monge constitui uma contribuição negativa ou positiva para a manutenção da autonomia (e.g.) do sistema social em que está incluso [Miłkowski 2015:5]<sup>142</sup>, assim como para a subsistência da linhagem ao longo do tempo. A conclusão poderá ser desfavorável no sentido em que o monge age de acordo com crenças que não contribuem para a preservação nem da sua autonomia, nem da autonomia do(s) sistema(s) social(is) em que participa<sup>143</sup>.

---

<sup>141</sup> António Zilhão diz-nos que este contraexemplo foi originalmente colocado à teoria do gene egoísta de Dawkins [1976].

<sup>142</sup> Como nos diz Miłkowski, esta mesma perspectivação ajudará também à consideração de actos heroicos (entre outros), talvez permitindo explicar «why my belief about philosophers causes me to write this paper instead of getting something to eat even if I'm hungry right now.» [2015:5].

<sup>143</sup> Existem outras objecções recentes ao projecto das semânticas de sucesso que não explorámos aqui [e.g. Nanay 2013 e de Prado Salas 2018].

Relativamente aos últimos dois contra-exemplos, (h) e (i), gostaríamos de apontar – e agradeceremos a António Zilhão por sublinhar este ponto – que se tratam de situações normativas ou morais, e que poderá não ser viável estender o princípio de Ramsey a crenças não-descriptivas, mas prescritivas. Nós entendemos que esta questão irá depender, em última instância, da possibilidade de naturalizar a Ética (e em particular da possível naturalização das normas e dos valores). A este respeito, subscrevemos quase por inteiro a teoria de Sterelny & Fraser [2016], segundo a qual verdades morais dependem ontologicamente (are grounded in) de factos acerca da cooperação.



## IV. APLICAÇÃO DA TESE DO TROPISMO AUTONÓMICO PARA A VERDADE À EPISTEMOLOGIA

Na secção III vimos como, de um modo geral, a tese do tropismo autonómico para a verdade permite fundamentar uma semântica de sucesso e explicar as dinâmicas de sucesso de representações particulares. Nesta secção iremos pensar a aplicação da tese do tropismo autonómico para a verdade a uma epistemologia naturalista.

### 1. O Corolário Evolucionário do Tropismo Autonómico para a Verdade: a Ponte Mílvia

É possível derivar um corolário relevante para a teoria evolutiva a partir da tese do tropismo autonómico para a verdade. Se há uma relação interna entre ser portador de conteúdos verdadeiros e ter sucesso interactivo, então a evolução natural irá seleccionar aqueles indivíduos que são portadores de conteúdos verdadeiros. Seguindo a terminologia de Paul Griffiths & John Wilkins [2015:203], chamaremos “ponte Mílvia” (*Milvian bridge*)<sup>144</sup> a este corolário que estabelece uma relação entre crenças verdadeiras e sucesso interactivo evolucionário:

Disposições com conteúdos verdadeiros são interna ou externamente, positiva e causalmente seleccionadas por preservarem a autonomia dos sistemas em que são instanciadas. Por oposição, disposições com conteúdos falsos são seleccionadas negativamente por contribuir para o equilíbrio termodinâmico dos seus portadores.

Este corolário permite-nos averiguar de um modo mais directo quais as consequências da tese do tropismo autonómico para a verdade para uma *epistemologia naturalizada formatada pela teoria evolutiva*, i.e. uma *epistemologia evolucionária* [exemplos tradicionais: Lorenz 1973, Campbell 1974, Vollmer 1975, Riedl 1984, Plotkin 1987, Wuketits 1990, Hooker 1995].

### 2. Argumentos Evolucionários a Favor e Contra a Fiabilidade da Cognição

Uma epistemologia evolucionária parte do pressuposto de que o «conhecimento é um problema em biologia evolucionária»<sup>145</sup> [Plotkin 1987:297] e que, enquanto tal, a teoria evolucionária tem algo a dizer acerca da nossa condição epistémica. Neste capítulo apresentaremos um

---

<sup>144</sup> A expressão deve-se a um evento histórico no qual Constantino terá vencido “a batalha da ponte Mílvia” porque a moral dos seus soldados era movida pela crença na verdade dos seus valores cristãos [*ibidem*].

<sup>145</sup> «Knowledge is a problem in evolutionary biology».

conjunto de argumentos que assume este pressuposto como fundamental para o problema da *justificação* das nossas crenças no contexto das teorias *fiabilistas* (*reliabilist*)<sup>146</sup>.

De acordo com uma teoria fiabilista da justificação, um agente epistémico S está justificado a acreditar em *p* se, e somente se, a crença verdadeira de S em *p* em *t* é gerada por um processo conducente à verdade [cf. Goldman & Beddor 2015 para uma problematização]. Uma teoria diz-se fiabilista porque diz dos processos que conduzem à verdade que são fiáveis; estas teorias são externistas porque sustentam que a fiabilidade dum processo é independente do possível acesso que o agente epistémico tenha às razões pelas quais o processo produz (em média) crenças verdadeiras.

A questão que nos interessa neste capítulo, porém, é a de saber se a selecção natural é um processo (causal) fiável, ou seja, se a selecção natural nos confere mecanismos cognitivos produtores de representações verdadeiras (com especial ênfase nas nossas crenças [cf. §II.2.]). Se sim, então a selecção natural é um fiador epistémico (*epistemic warrantor*) capaz de justificar as nossas crenças; se não, então a selecção natural é um derrotador epistémico (*epistemic defeater*) das nossas crenças<sup>147</sup>.

Para que possamos averiguar esta tese será primeiramente necessário discernir entre dois tipos de fiabilidade. O primeiro, já introduzido acima, e a que poderemos chamar “*fiabilidade epistémica*” (*truth-reliability*), diz respeito à questão de saber se um dado processo conduz ou não à verdade. O segundo, a que poderemos chamar “*fiabilidade adaptativa*” (*fitness-reliability*), diz respeito à questão de saber se um dado processo é ou não adaptativo. A teoria evolucionária descreve o processo de selecção natural como maximizador da fiabilidade adaptativa. Se este é também um processo que maximiza a fiabilidade epistémica é uma questão ainda hoje largamente discutida, mas relativamente à qual responderemos afirmativamente, em acordo com o corolário da ponte Mílvia. Apresentaremos então “argumentos evolucionários” epistemicamente *optimistas* [AEO], contra os quais apresentaremos outros *cépticos* [AEC], e outros *pesimistas* [AEP].

Em geral, todos os argumentos evolucionários partilham uma estrutura simplificada conforme à seguinte [a formulação é de Kahane 2011:106]:

1. A crença de S em *p* é explicada por X. (Premissa *causal*).
  2. X (não) é um processo fiável. (Premissa *epistémica*).
- ∴ A crença de S em *p* (não) é justificada.

Atendendo a esta estrutura geral, repare-se que o facto de uma crença ter uma explicação causal é em si bastante inócuo para o problema da justificação, pois «o mero facto de que há uma explicação causal para uma crença não afecta a sua justificação»<sup>148</sup> [Kahane 2011:105]. Do mesmo modo, «o facto de que as faculdades cognitivas humanas têm origens evolutivas não providencia razões para aceitar a reivindicação de que elas são fiáveis com respeito a gerarem

---

<sup>146</sup> Outras versões de epistemologia evolucionária e fiabilista existem que se preocupam, p.ex., com a própria definição do que é o conhecimento. A título de curiosidade, terá sido também Ramsey [1931] quem primeiramente sugeriu uma teoria fiabilista do conhecimento.

<sup>147</sup> Plantinga [1993] defende que este projecto naturalista é auto-refutante. A este respeito, considere-se a resposta de Fitelson & Sober [1997].

<sup>148</sup> «the mere fact that there is a causal explanation of a belief does nothing to affect its justification»

crenças verdadeiras»<sup>149</sup> [Sage 2004:95]. Não devemos, pois, confundir razões com causas, de modo a evitar cometer uma falácia genética.

Repare-se também que estes argumentos têm por alvo a justificação da nossa atitude epistémica relativamente a representações produzidas por um determinado processo. Porém, estes argumentos nada nos dizem relativamente à nossa atitude face a representações *específicas* produzidas por esse determinado processo. Assim, estes argumentos só nos podem dizer se estamos justificados a acreditar nas representações produzidas por X, mas não nos permitem concluir relativamente a uma determinada representação *p* produzida por X se ela é verdadeira ou falsa<sup>150</sup>.

Note-se ainda que os argumentos evolucionários acerca da justificação epistémica são *circulares*, pois a verdade da sua conclusão depende da fiabilidade epistémica do mesmo processo causal que explica a crença na sua conclusão. Porém, a sua circularidade é *informativa*. É informativa ou *não viciosa* porque os argumentos não são *cognitivamente impotentes* ou «patologicamente ineficazes contra quem não aceitar já as suas conclusões»<sup>151</sup> [Lipton 2000:189; cf. ainda Lipton 1991]<sup>152</sup>. De facto, parece-nos que Stathis Psillos [1999:84] está correcto na sua observação de que a viciosidade de um argumento circular depende da própria teoria da justificação empregue porque, como o fiabilismo é uma teoria da justificação *externista* – querendo com isto dizer que os operadores da justificação são não-mentais ou inconscientes –, a justificação é independente de termos ou não *acesso* (directo) às *razões* pelas quais estamos justificados em acreditar nas representações produzidas por um determinado processo [cf. Pappas 2017]. Assim, a conclusão de um argumento evolucionário pode acrescentar àquilo que já sabemos, concluindo que um mecanismo cognitivo (não) é fiável, asserção essa que será inteiramente dependente da fiabilidade epistémica do mecanismo cognitivo (e.g. das regras de inferência produzirem, em média, crenças verdadeiras) e não de qualquer acesso que possamos ter às razões pelas quais esse mecanismo cognitivo é fiável. Relevantemente, a fiabilidade é avaliada *externamente*, enquanto «equivale a uma hipótese física de que existe uma correlação entre o *output* de um método [no nosso caso, um mecanismo cognitivo] e o estado do mundo»<sup>153</sup> [Lipton 2000:180]. A epistemologia naturalizada de Quine já dava conta destas subtilidades quando este nos disse que:

«Não estou a apelar à biologia Darwiniana para justificar a indução. Isto seria circular, uma vez que o conhecimento biológico depende da indução. Em vez disso, estou a pressupor a eficácia da indução, e depois a observar que a biologia Darwiniana, se verdadeira, nos ajuda a explicar porque é que a indução é tão eficaz quanto é»<sup>154</sup> [Quine 1975:50]

---

<sup>149</sup> «the fact that human cognitive faculties have an evolutionary origin provides no reason to accept the claim that they are reliable with respect to generating true beliefs.»

<sup>150</sup> Neste tanto concordamos com Kahane [2011:108], mas verificamos que isto tem por consequência que podemos estar injustificados a acreditar em crenças que não obstante são verdadeiras. Contra Kahane, contudo, parece-nos ser o caso que, estando justificados a acreditar nos produtos de X, então é também o caso que as representações produzidas por X serão *provavelmente* verdadeiras.

<sup>151</sup> «pathologically ineffective against anyone who did not already accept its conclusion.»

<sup>152</sup> Pense-se no exemplo de David Hume relativamente à justificação indutiva da indução: se uma indução não é justificada, tanto pior será o estatuto epistémico de uma justificação indutiva da indução.

<sup>153</sup> «tantamount to a physical hypothesis that there is a correlation between the output of the method and the state of the world.»

<sup>154</sup> «I am not appealing to Darwinian biology to justify induction. This would be circular, since biological knowledge depends on induction. Rather, I am granting the efficacy of induction, and then observing that Darwinian biology, if true, helps explain why induction is as efficacious as it is.»

Posto isto, o peso dos argumentos evolucionários encontra-se então na premissa epistémica. É necessário investigar se o processo causal da selecção natural faz convergir a fiabilidade adaptativa e a fiabilidade epistémica, *i.e.* temos de investigar a hipótese empírica da selecção natural ser um fiador epistémico (com menção a operadores externos à racionalidade dos agentes).

Face a esta hipótese de investigação há que recordar que *a evolução natural é um processo causal ateleológico e que não tende, portanto, para a verdade*. Como nos diz Michael Ruse, a evolução «vai a lado nenhum – e de forma bastante lenta»<sup>155</sup> [1986:203]. Noutras palavras, não pretendemos defender aqui que a evolução natural *per se* é um fiador epistémico. A evolução natural não garante sequer a existência de seres vivos amanhã, e muito menos garantirá a existência de seres vivos com um extraordinário leque de crenças verdadeiras. De facto, se atendermos às fundações termodinâmicas do processo evolutivo, tudo indica que num futuro longínquo deixará de existir a energia ordenada livre da qual a manutenção da autonomia dos biosistemas depende<sup>156</sup>. Consequentemente, o nosso argumento dirá unicamente respeito à fiabilidade epistémica da selecção natural, e não à fiabilidade epistémica da evolução como um todo.

A hipótese empírica da selecção natural ser um fiador epistémico segue-se da ponte Mílvia. Manifestamente, a tese do tropismo autonómico para a verdade diz respeito a conteúdos, mas o corolário evolucionário generaliza a tese a qualquer produtor de conteúdos. A ideia é simples: se crenças verdadeiras são evolucionariamente vantajosas, a sua eficácia causal será responsável pela selecção indirecta dos seus produtores (os mecanismos cognitivos). Mecanismos cognitivos que produzam crenças verdadeiras são então selecionados devido ao contributo positivo dos seus produtos para a manutenção da autonomia. Por consequência, a premissa epistémica do nosso argumento evolucionário optimista será a seguinte:

- A. Só há selecção causal positiva de mecanismos cognitivos quando há um tropismo para a verdade.

Esta hipótese A é em geral considerada como sendo extremamente radical, sendo, na maioria das vezes, ignorada ou rejeitada [e.g. Simpson 1963, Quine 1969, Sober 1981, Fodor 1981, Dennett 1987, Feldman 1988, Stich 1990, Plantinga 1996, Fales 1996, Sage 2004, e de Cruz *et al.* 2011]. Mas há duas excepções [nomeadamente, Griffiths & Wilkins 2015, e Michael Vlerick & Alex Broadbent 2015], lado-a-lado das quais firmaremos a nossa defesa desta hipótese.

Opõe-se a A uma hipótese B mais moderada:

- B. O valor esperado da fiabilidade adaptativa dos mecanismos cognitivos com tropismos para a verdade é superior ao valor esperado da fiabilidade adaptativa dos mecanismos cognitivos sem tropismo para a verdade.

De acordo com os defensores de qualquer destas duas hipóteses, a selecção natural desempenha o papel (mais ou menos fraco) de fiador epistémico:

---

<sup>155</sup> «is going nowhere – and rather slowly at that»

<sup>156</sup> Num breve exercício de especulação diríamos que a termodinâmica em si também não poderá ser um fiador epistémico, pois a teoria diz-nos que a complexidade do Cosmos irá eventualmente reduzir-se devido à ausência da energia ordenada necessária à realização de trabalho, e como os constrangimentos económicos à manutenção de sistemas cognitivos aumentarão, a ausência de complexidade ambiental deixará de justificar a manutenção de sistemas cognitivos tão custosos, de complexos que são.

«As nossas percepções dão-nos realmente representações verdadeiras, ainda que incompletas, do mundo exterior porque isso foi e é uma necessidade biológica posta em nós pela selecção natural; se não fosse assim, não estaríamos aqui.»<sup>157</sup> [Simpson 1963:98]

«A selecção natural garante que a maior parte das crenças de um organismo serão verdadeiras e a maioria das suas estratégias racionais.»<sup>158</sup> [Dennett 1987:75]<sup>159</sup>

E isto porque:

«A selecção Darwiniana garante que os organismos ou conhecem os fundamentos da lógica ou tornam-se póstumos.»<sup>160</sup> [Fodor 1981:121]<sup>161</sup>

«O macaco que não teve uma percepção realista do ramo da árvore para o qual saltou tornou-se rapidamente um macaco morto – e, como tal, não se tornou um dos nossos antepassados.»<sup>162</sup> [Simpson 1963:98]

E novamente:

«Criaturas inveteradamente erradas nas suas induções têm uma tendência patética mas louvável para se extinguirem antes de reproduzirem o seu género.»<sup>163</sup> [Quine 1969:126]

E para concluir com um exemplo mais detalhado:

«[S]e um ser tem crenças, é melhor (isto é, mais propício para a sobrevivência) que ele tenha crenças verdadeiras do que crenças falsas. Crenças verdadeiras acerca de onde está a nossa comida são mais úteis para encontrar comida, e sobreviver, do que crenças falsas. Semelhantemente, crenças verdadeiras acerca de onde estão os nossos predadores e acerca de como lhes escapar contribuem mais para aumentar a nossa sobrevivência do que crenças falsas acerca destes assuntos. Então, é mais provável a selecção natural seleccionar crentes que têm sobretudo crenças verdadeiras. A melhor forma, talvez a única forma, de crentes terem sobretudo crenças verdadeiras é terem mecanismos ou estratégias formadoras de crenças fiáveis. Mecanismos ou estratégias fiáveis são aquelas que levam sobretudo a crenças verdadeiras. Assim, a selecção natural irá seleccionar crentes que têm mecanismos formadores de crenças fiáveis.»<sup>164</sup> [Feldman 1988:218]

---

<sup>157</sup> «Our perceptions do give true, even though not complete, representations of the outer world because that was and is a biological necessity, built into us by natural selection if it were not so, we would not be here.»

<sup>158</sup> «Natural selection guarantees that most of an organism's beliefs will be true, most of its strategies rational»

<sup>159</sup> O argumento de Dennett é inspirado num princípio de caridade Davidsoniano segundo o qual a maioria das nossas crenças têm de ser verdadeiras para que haja racionalidade, e de acordo com o qual uma comunicação linguística de sucesso pressupõe que há um corpo partilhado de crenças [cf. Davidson 1977].

<sup>160</sup> «Darwinian selection guarantees that organisms either know the elements of logic or become posthumous.»

<sup>161</sup> Mais tarde, Fodor veio a alterar a sua posição [cf. Fodor & Piatelli-Palmarini [2010].

<sup>162</sup> «The monkey who did not have a realistic perception of the tree branch it jumped for was soon a dead monkey – and, therefore, did not become one of our ancestors.»

<sup>163</sup> «Creatures inveterately wrong in their inductions have a pathetic but praiseworthy tendency to die out before reproducing their kind.»

<sup>164</sup> «[I]f a being has beliefs at all, it is better (that is, more conducive to survival) for it to have true beliefs than false beliefs. True beliefs about where one's food is are more helpful for finding food, and surviving, than are false beliefs. Similarly, true beliefs about where one's predators are and how to escape them are more survival enhancing than false beliefs about these matters. So, natural selection is likely to select for believers that have mostly true beliefs. The best way, perhaps the only way, for believers to have mostly

A diferença relevante entre A e B é que, de acordo com B, a selecção natural pode ser um derrotador epistémico, ainda que haja uma correlação forte entre fiabilidade adaptativa e fiabilidade epistémica. E isto deve-se a uma de duas razões. A formulação da primeira deve-se a James Sage [2004:98], que nos diz que nem todas as características dos organismos que se podem dizer adaptativamente fiáveis se podem também dizer epistemicamente fiáveis. Ele dá-nos o exemplo de traços *não-cognitivos* que podem ser fiáveis adaptativamente e que *prima facie* não serão fiáveis epistemicamente (o que ele conclui perante a observação de que os produtos de disposições não-cognitivas não são portadores de verdade). Contudo, de acordo com o Interactivismo [cf. §II.], características não-cognitivas serão, ainda assim, portadoras de conteúdos *primitivos*, pelo que este contraexemplo de Sage não se repercute negativamente na nossa perspectiva, uma vez que esta nos permite tomar a fiabilidade adaptativa a par da existência de tropismos para a verdade para qualquer característica que for uma aptação. Uma segunda razão frequentemente avançada a favor de B deve-se à possibilidade de haver faculdades *cognitivas* adaptativamente fiáveis que, não obstante, também não serão epistemicamente fiáveis. Dizem-nos estes autores que:

«[N]ão podemos deduzir a capacidade de “produzir-verdade” de um traço cognitivo a partir do seu valor para a sobrevivência.»<sup>165</sup> [Thomson 1995:173]<sup>166</sup>

E, mais explicitamente com ênfase nos mecanismos cognitivos responsáveis pela produção de teorias científicas:

«Não há garantia para passar de “estratégia cognitiva x não resulta num decrescimento fatal na fitness” para “estratégia cognitiva x confere-nos teorias aproximadamente verdadeiras»<sup>167</sup> [Thomson 1995:174]

Ou ainda:

«[S]e as nossas faculdades cognitivas foram originadas como Dawkins pensa [*i.e.* por selecção natural], então o seu último propósito ou função (...) será qualquer coisa como sobrevivência (de indivíduos, espécies, gene ou genótipo); mas parece inicialmente duvidoso que entre as suas funções – últimas, próximas ou outras – esteja a produção de crenças verdadeiras»<sup>168</sup> [Plantinga 1993:218]

Contra este pessimismo e *sensu* a tese do tropismo autonómico para a verdade e seu corolário evolucionário, argumentaremos que, sim, é possível deduzir o valor de verdade de uma representação a partir da sua contribuição sistemática para a sobrevivência do(s) seu(s) portador(es). Todavia, um argumento evolucionário optimista mais fraco, por via da hipótese B, precisaria apenas de demonstrar que «no todo e a longo prazo, organismos serão mais fit – irão

---

true beliefs is for them to have reliable belief-forming mechanisms or strategies. Reliable mechanisms or strategies are ones that leads mostly to true beliefs. Hence, natural selection will select believers that have reliable belief-forming mechanisms.»

<sup>165</sup> «[O]ne cannot deduce the “truth-producing” capacity of a cognitive trait from its survival value.»

<sup>166</sup> Paul Thomson refere-se à recensão crítica elaborada por Bruce Hauptli [1994] ao livro de Nicholas Rescher [1989].

<sup>167</sup> «There is no warrant for moving from “cognitive strategy x does not result in a fatal decrease in fitness” to “cognitive strategy x delivers approximately true theories”.»

<sup>168</sup> «[I]f our cognitive faculties have originated as Dawkins thinks [*i.e.* through natural selection], then their ultimate purpose or function (...) will be something like *survival* (of individuals, species, gene or genotype); but then it seems initially doubtful that among their functions – ultimate, proximate or otherwise – would be the production of true beliefs.»

superar a concorrência dos seus conspécíficos – se o seu rácio de crenças verdadeiras para falsas for mais elevado»<sup>169</sup> [Stich 1990:59].

A estas duas hipóteses optimistas podemos ainda opor uma *céptica*:

- C. O tropismo para a verdade dos mecanismos cognitivos é independente da sua fitness.

E outra *pessimista*:

- D. O valor esperado da fiabilidade adaptativa dos mecanismos cognitivos com tropismos para a verdade é inferior ao valor esperado da fiabilidade adaptativa dos mecanismos cognitivos sem tropismo para a verdade.

Concluimos assim com quatro hipóteses passíveis de ocupar o lugar da premissa epistémica dos argumentos evolucionários, a serem descortinadas por investigação empírica. Não atenderemos à hipótese C, mas sim às restantes<sup>170</sup>. A hipótese A é a mais frágil, pois é também a mais forte, negando que haja quaisquer crenças falsas que possam sistematicamente contribuir para a fitness dos seus portadores. A hipótese B considera que há mecanismos cognitivos produtores de crenças falsas que são adaptativamente fiáveis, mas que «o facto de que existem crenças falsas particulares que aumentam a fitness inclusiva de um organismo não desafia seriamente [B]»<sup>171</sup> [Sage 2004:102].

Em todo o caso, qualquer proponente de um argumento evolucionário concordará que as nossas representações, sejam elas verdadeiras ou falsas, são relevantes para a nossa fitness<sup>172</sup>. Isto porque, se as nossas representações determinam o nosso comportamento no mundo, então a existência continuada de biosistemas com mecanismos cognitivos significa que estes contribuem para a fiabilidade adaptativa dos seus portadores. Ou isso ou então todo o comportamento é praticamente neutro para a fitness ou negligenciavelmente deletério, o que é uma hipótese dúbia se considerarmos que a aquisição de crenças pode ser energeticamente dispendiosa (tome-se p.ex. os trabalhos necessários ao desenvolvimento de algumas representações, como a própria teoria evolucionária contemporânea, cujos trâmites dependem dos esforços cognitivos de muitos agentes ao longo de centenas de anos). Além do mais, e de um modo geral, o nosso sistema nervoso apresenta uma taxa metabólica de repouso bastante alta (e superior ainda com a realização de actividade física) [cf. Isler & van Schaik 2006]. Assim, *independente das propriedades dos nossos mecanismos cognitivos*, eles devem contribuir para a fitness

---

<sup>169</sup> «All the advocate of the current argument need claim is that on the whole and in the long run, organisms will be more fit – they will outcompete their conspecifics – if their ratio of true beliefs to false ones is higher.»

<sup>170</sup> Talvez a hipótese C possa ser associada à ideia fundamental de que «a função fundamental da cognição e as atitudes intencionais fundamentais envolvidas na actividade cognitiva podem ser caracterizadas sem apelo ao conceito de verdade» («the fundamental function of cognition and the fundamental intentional attitudes involved in cognitive activity can be characterized without appeal to the concept of truth») [Kvanvig 2005:293].

<sup>171</sup> «the fact that there are particular false beliefs that increase an organism's inclusive fitness does not seriously call [B] into question.»

<sup>172</sup> Plantinga [1993] acredita que crenças metafísicas (que ele exemplifica com a crença no naturalismo) são neutras para a fitness, o que é obviamente falso se considerarmos o sucesso das acções protagonizadas por um sismólogo naturalista durante um terramoto comparativamente às de um indivíduo anti-naturalista que acredita que realizar um ritual religioso irá aumentar a sua probabilidade de sobreviver ao terramoto.

se o seu custo energético tão elevado é suportado pela evolução por selecção natural, e os epistemólogos evolucionários parecem estar bem cientes disto, quando nos dizem que:

«O *Homo sapiens* especializou-se, mais do que qualquer outra espécie, na inteligência como uma estratégia de sobrevivência. (...) O nosso grande investimento em encéfalos grandes e em corpos noutros aspectos medíocres torna ainda mais improvável que recursos fossem despendidos em mecanismos complexos de formação e processamento de crenças sem utilidade prática.»<sup>173</sup> [Fales 1996:440]

«As hipóteses de que a crença não tem efeito no comportamento, ou que ter crenças reduz a fitness, são completamente absurdas. Se este fosse o caso, então haveria forte selecção para não ter crenças, dado que as crenças são caras de adquirir e manter.»<sup>174</sup> [Griffiths & Wilkins 2015:235].

Tendo feito esta introdução, podemos então construir os seguintes argumentos evolucionários (substancialmente inspirados nas construções silogísticas de Sage [2004:100] e de Helen de Cruz *et al.* [2011:519]):

AEO		AEC	AEP
1. Os biosistemas são autónomos.			
2. Representações são estados funcionais dos organismos que implementam ou incorporam o seu endosso de um estado particular de coisas como actual e são, enquanto tal, guias para interacções com condições ambientais ou interiores esperadas.			
3. Em biosistemas cujos mecanismos cognitivos produzem representações (não-mínimas), estas contribuem negativa ou positivamente para a fitness (não são neutras para a fitness).			
4. A selecção natural favorece características que aumentam a fitness dos seus portadores <sup>175</sup> .			
5. A. Só há selecção causal positiva de mecanismos cognitivos quando há tropismos para a verdade.	B. O valor esperado da fiabilidade adaptativa dos mecanismos cognitivos com tropismo para a verdade é superior ao valor esperado da fiabilidade adaptativa dos mecanismos cognitivos sem tropismo para a verdade.	C. O tropismo para a verdade dos mecanismos cognitivos é independente da sua fitness.	D. O valor esperado da fiabilidade adaptativa dos mecanismos cognitivos tropismos para a verdade é inferior ao valor esperado da fiabilidade adaptativa dos mecanismos cognitivos sem tropismo para a verdade.
6. A selecção natural não é o único impulso/força/mecanismo da evolução <sup>176</sup> .			
7. Os mecanismos evolutivos para além da selecção natural não têm uma relação nomológica com aumentos positivos para a fitness.			

<sup>173</sup> «*Homo sapiens* has, more than any other species, specialized in intelligence as a survival strategy (...) Our heavy investment in big brains and otherwise mediocre bodies makes it all the more unlikely that resources would be wasted on elaborate belief-forming and processing mechanisms that have no practical utility.»

<sup>174</sup> «The hypotheses that belief has no effect on behaviour, or that having beliefs reduces fitness, are completely absurd. If this were the case then there would be strong selection for not having beliefs, as beliefs are very costly to acquire and maintain.»

<sup>175</sup> Com a ressalva de que existem fenómenos selectivos não-adaptativos [cf. Anexo B e §IV.3.].

<sup>176</sup> E.g.: mutações (aleatórias ou não), deriva genética e fluxo génico (por migração) [cf. Anexo A].



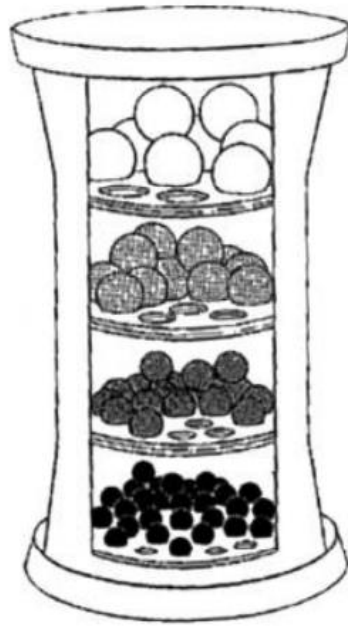
8. A selecção natural é o principal determinante da evolução dos sistemas cognitivos.		8*. A selecção natural não é o principal determinante da evolução dos sistemas cognitivos.	
9. Existem constrangimentos lógicos, económicos e ambientais inerentes à selecção natural de tropismos para verdade.			
10. Os mecanismos cognitivos humanos são, com uma grande probabilidade, epistemicamente fiáveis.	10*. Os mecanismos cognitivos humanos são, com uma probabilidade de pelo menos acima de 50%, epistemicamente fiáveis.	10**. Nada podemos concluir acerca da fiabilidade epistémica dos mecanismos cognitivos humanos.	10***. Os mecanismos cognitivos humanos são, com uma probabilidade inferior a 50%, epistemicamente fiáveis.

Neste ensaio iremos defender um argumento com a forma 1-5A, *suspendendo o juízo relativamente a 8-10 ou 10\**. Isto é, iremos defender que a selecção natural é um fiador epistémico, mas que não sabemos o quão influentes são os determinantes evolutivos para além da selecção natural. Porque será especialmente relevante a veracidade da nossa hipótese A, começaremos por clarificá-la; de seguida atenderemos à premissa 9; e concluímos a secção com uma análise de contraexemplos. Na secção conclusiva deste ensaio atenderemos brevemente às premissas 8 e 8\*, de cujas conclusões 10 ou 10\* dependem, e, por sua vez, depende também a possibilidade de uma teoria naturalista e realista do sucesso da Ciência.

### 3. A Distinção entre “Selecção *Simpliciter*” e “Selecção Causal”

Para clarificar a natureza da hipótese A, que só há selecção causal positiva de mecanismos cognitivos quando há tropismos para a verdade, começamos por introduzir uma distinção que nos ajudará a compreender a própria natureza da selecção natural. A distinção, como entendida por Sober [1993a], é entre “selecção de” e “selecção devido a” (*selection of* e *selection for*, respectivamente).

Sober convida-nos a pensar num “brinquedo de selecção”. O brinquedo tem vários níveis demarcados por separadores com buracos circulares de um mesmo tamanho. De cima, para baixo, a cada nível que se segue os buracos têm dimensões cada vez menores. Dentro do brinquedo há também uma série de bolas com diversos tamanhos e cada tamanho apresenta uma mesma cor. As bolas começam todas no nível superior e, conforme o brinquedo é abanado, vão progressivamente sendo filtradas, nível após nível, até que se apresentem conforme a imagem:



Uma réplica do “brinquedo de selecção” após selecção [Sober 1993a:99].

Com base neste brinquedo, Sober propõe então a seguinte distinção: a cada nível há *selecção de bolas* de um determinado tamanho e de uma determinada cor; porém, só há *selecção devido ao tamanho das bolas* e não devido à sua cor. Há, pois, *selecção de objectos* (bolas) *devido às suas propriedades* (o seu tamanho). Sober explicita:

«“Selecção de” diz respeito aos *efeitos* de um processo de selecção, enquanto “selecção devido a” descreve as suas *causas*. Dizer que há selecção devido a uma dada propriedade significa que ter aquela propriedade causa sucesso na sobrevivência e reprodução. Mas dizer que um dado tipo de objecto foi selecionado é dizer meramente que o resultado de um processo de selecção foi o aumento da representação daquele tipo de objecto. (...) “Selecção devido a” é o conceito causal por excelência.»<sup>177</sup> [1993a:100]

Dada a natureza causal da selecção de propriedades, é-nos possível analisar contrafactualmente as propriedades dos objectos que são causalmente selecionadas e distingui-las daquelas que, como a cor das bolas no brinquedo, são selecionadas colateralmente ou irrespectivamente do seu contributo para a fitness. Considerem-se então alguns exemplos à luz desta distinção.

«Por exemplo, o casaco grosso do urso polar é selecionado pelas suas propriedades térmicas, mas um casaco quente também tem de ser um casaco pesado, de tal forma que o peso é tão fiavelmente selecionado como o quente. (...) Se o casaco do urso polar fosse quente mas não pesado, teria ainda assim sido selecionado. Mas se o casaco fosse pesado e não quente, não teria sido selecionado. Assim, o calor, e não o peso, é o alvo da selecção.»<sup>178</sup> [Goode & Griffiths 1995:103]

<sup>177</sup> «“Selection of” pertains to the *effects* of a selection process, whereas “selection for” describes its *causes*. To say that there is selection for a given property means that having that property causes success in survival and reproduction. But to say that a given sort of object was selected is merely to say that the result of the selection process was to increase the representation of that kind of object. (...) “Selection for” is the causal concept *par excellence*.»

<sup>178</sup> «For example, the thick coat of a polar bear is selected for its thermal properties, but a warm coat must also be a heavy coat, so heaviness is selected just as reliably as warmth. (...) *If the polar bear's coat had*

*Mutatis mutandis* para qualquer fenómeno selectivo não-adaptacionista [cf. Anexo B]. A título de exemplo, sabemos que o vermelho do sangue é dado pela relação entre a hemoglobina e o dioxigénio devido ao seu Fe<sub>4</sub>, mas que não é por isso que existe selecção de sangue que (por acaso) é vermelho. A hemoglobina carrega o oxigénio de que o nosso corpo necessita, e é por isso que é causalmente seleccionada. O vermelho é apenas um *subproduto* da selecção. Brevemente, fenómenos de selecção não-adaptacionista deste tipo são *tolerados* porque os seus efeitos não-adaptativos são: (1) negligenciáveis, (2) neutros, ou (3) compensados pelos efeitos positivos das propriedades causalmente seleccionadas (que podem ter benefícios superiores aos subprodutos, atenuar o efeito destes, ou até mesmo bloqueá-los).

Assim, tendo por base esta distinção, quando dizemos que só há selecção causal positiva de mecanismos cognitivos quando estes têm tropismos para a verdade, queremos com isto dizer que só estes mecanismos cognitivos exibem propriedades causais que aumentam eficazmente a fitness dos seus portadores, o que se segue da tese de que os conteúdos verdadeiros têm propriedades causais internamente relacionadas com o ambiente. Porém, também *existe selecção de mecanismos cognitivos que produzem crenças falsas*; mas, à semelhança do que sucede com fenómenos selectivos não-adaptacionistas, isto dá-se somente quando os efeitos nefastos das crenças falsas são toleráveis relativamente à satisfação da nossa autonomia.

A lição a reter é a de que fenómenos selectivos não-adaptacionistas dão conta da limitação da selecção natural *simpliciter* enquanto fiador epistémico: a selecção natural tolera algumas disposições com valores potencialmente negativos para a fitness; enquanto tal não é um fiador epistémico 100% fiável [Stich 1990:66-7 e Kitcher 1985:215 sublinham muito bem este facto] (a percentagem da fiabilidade dos nossos mecanismos cognitivos, porém, é mais uma vez deixada em aberto).

#### 4. A Complementaridade das Explicações com Base na Fitness e com Base em Tropismos para a Verdade

Como temos vindo a argumentar, o nosso fundo teórico interactivista permite-nos integrar o sucesso semântico sistemático no sucesso evolucionário sistemático. O sucesso semântico sistemático é a forma como se manifesta ao nível cognitivo o sucesso epistémico sistemático. Estas correspondências, dum ponto de vista da explicação do sucesso das acções, requerem que pensem as explicações evolucionárias, que têm por base a selecção natural (causal), e as explicações epistémicas, que têm por base um tropismo para a verdade, como *complementares*, e não como alternativas antagónicas. Neste capítulo propomo-nos desenvolver o sentido desta complementaridade.

Comecemos por observar que as consequências desta complementaridade são múltiplas. A título de exemplo: se estas explicações forem de facto complementares, então torna-se imediata a explicação de porque é que nenhum mecanismo produtor de crenças falsas contribui sistematicamente para aumentos da fitness dos seus portadores – tal seria logicamente impossível. Tais mecanismos produtores de crenças falsas, a existirem, teriam de ser adaptativamente fiáveis, passando o teste contrafactual que demonstraria que teriam sido causalmente seleccionados (e não seleccionados *simpliciter*).

---

*been warm but not heavy, it would still have been selected. But if the coat had been heavy but not warm, it would not have been selected. So warmth, rather than weight, is the target of selection.»*

Propomos acompanhar a análise contrafactual de Griffiths & Wilkins [2015] para melhor compreender de que modo é que estas explicações são complementares. Esta análise parte do debate em torno da chamada “indeterminação das funções biológicas” [cf. e.g. Goode & Griffiths 1995 e Neander 1995], que tem origem na observação de que por vezes parece ser impossível aferir, com base em toda a evidência disponível, qual, de entre várias descrições históricas da evolução de uma determinada característica, é a verdadeira descrição do seu processo selectivo causal. Isto é, por vezes uma mesma característica é passível de várias histórias selectivas (*just-so-stories* [Gould & Lewontin 1979], *pace* Kipling) igualmente legítimas. E isto é sobremaneira problemático para a atribuição de funções porquanto, no contexto da teoria etiológica [cf. §1.3-4.], conhecer a história evolutiva exacta de uma característica é uma condição necessária para a atribuição de funções. Na ausência de uma descrição exacta, a função permanecerá, enquanto tal, *indeterminada*.

O sistema visual dos sapos *Rana pipiens* permite-nos ilustrar este debate de uma forma exemplar. Passa-se que, após várias experiências, um grupo de cientistas [Lettvin *et al.* 1959:231] concluiu que várias pistas ambientais, todas elas caracterizáveis por um determinado tamanho e movimento, eram capazes de estimular um mesmo estado (mental) perceptual no qual os sapos reagem de modo a capturar alvos com as suas línguas preênsais. Concluíram então que os sapos foram selecionados para responderem a qualquer objecto que reunisse essas características (pequeno, escuro e com um certo movimento). Todavia, concluíram também que este comportamento não é 100% adaptativamente fiável, visto que leva os sapos a reagir na presença de *qualquer* objecto pequeno e escuro que se movimenta no seu campo de visão (exemplo: uma bolinha de papel preta suspensa num fio de marioneta, que não é nutritiva). Atendendo a estas experiências, vários biólogos e filósofos da biologia debateram qual seria a função deste mecanismo perceptivo. Goode & Griffiths [1995:100] documentam:

«Kim Sterelny (1990) argumentou que a função do mecanismo é detectar e capturar moscas (ou quaisquer que fossem as taxa de insectos que formavam a presa dos ancestrais do sapo). Lawrence A. Shapiro (1992) argumentou que a função do mecanismo é detectar comida. Finalmente, Karen Neander (1995) argumentou que a função do mecanismo é detectar coisas pequenas e escuras em movimento. Sapos ancestrais saíram-se bem porque conseguiram detectar coisas pequenas e escuras em movimento. Também se saíram bem porque conseguiam detectar moscas, e porque conseguiam detectar comida.»<sup>179</sup>

Ora, como as últimas frases da citação dão a entender, Goode & Griffiths sugerem que todas estas descrições do processo selectivo são na realidade compatíveis. Elas apenas diferem relativamente ao nível de abstracção em que as explicações decorrem. De facto, se empregarmos uma análise contrafactual destas explicações, a sua compatibilidade torna-se evidente. Se o sapo não reagisse a pequenas coisas escuras em movimento, então não apanharia moscas, que são a sua comida, o que contribuiria em fim para um decréscimo da sua fitness. Uma das explicações é neuroanatômica, outra é ecológica, dando conta das *causas diferenciais dos valores de fitness*, e outra decorre ao nível das *consequências para a fitness* (sendo independente do realizador do papel causal) [cf. Sober 1984]. Decorrendo estas explicações em diferentes níveis de abstracção, captando propriedades que podem ocorrer simultaneamente, elas não são derradeiramente alternativas em competição.

---

<sup>179</sup> «Kim Sterelny (1990) has argued that the function of the mechanism is to detect and capture flies (or whatever taxa of insects form the frog’s ancestral prey). Lawrence A. Shapiro (1992) has argued that the function of the mechanism is to detect food. Finally, Karen Neander (1995) has argued that the function of the mechanism is to detect small dark moving things. All three positions gain support from the etiological theory of proper function. Ancestral frogs did well because they could detect small dark moving things. They also did well because they could detect flies, and because they could detect food.»

O exemplo anterior do urso polar [§IV.3.] também nos permite dar conta desta complementaridade: o urso polar tinha um casaco quente que também era pesado, e concluímos que é por ser quente e não por ser pesado que o seu pelo foi (e é) selecionado. Ora, poderíamos dizer que a selecção do casaco quente é uma alternativa à selecção de um contribuidor positivo para a fitness? Seria absurdo responder que sim neste caso. *É somente quando estamos a falar acerca de um mesmo nível de selecção que podemos perguntar qual das propriedades foi selecionada por alternativa à outra* [Goode & Griffiths 1995:104]. Mais que isso, não podemos abdicar de nenhuma das explicações decorrentes nos diferentes níveis, pois «[s]upor que uma narrativa histórica particular não acrescenta nada a um amplo modelo ecológico é pressupor que uma explicação causal detalhada não acrescenta nada a uma idealização sugestiva»<sup>180</sup> [*idem*:107]. Do mesmo modo, também não podemos pressupor que um modelo ecológico não acrescenta nada a explicações que são cegas à natureza dos portadores de valores de fitness. «A classificação de um único domínio de objectos numa série de vocabulários cada vez mais abstractos permite aceder a generalizações que estão escondidas em níveis mais baixos pela múltipla realização dos tipos que entram nessas generalizações»<sup>181</sup> [*idem*:103]. Há, pois, que *integrar a pluralidade de explicações não-alternativas* para termos acesso a uma compreensão mais completa do sistema alvo [cf. Darden & Maull 1977, Mitchell 2003 & 2004:84-6].

É com base neste racional que Griffiths & Wilkins [2015] sugerem analisar a complementaridade das explicações do sucesso que têm por base a selecção natural (causal) e aquelas outras, epistémicas, que têm por base tropismos para a verdade: elas são complementares («aumentos na fitness são explicados pelo sucesso da detecção de verdade e reduções na fitness pelo insucesso na detecção da verdade»<sup>182</sup> [Griffiths & Wilkins 2015:239]). Assim, mecanismos cognitivos contribuem positivamente para a fitness *porque* têm tropismos para a verdade, ou, então, têm tropismos para a verdade *de modo a* aumentar a fitness, sendo que a verdade desta complementaridade se pode aferir pelo emprego de uma análise contrafactual, bastando para isso perguntar: se os mecanismos não tivessem um tropismo para a verdade, a evolução selectiva dos mecanismos cognitivos teria decorrido da mesma maneira? A resposta de Griffiths & Wilkins é que não, e é deste modo que propõem a hipótese A, que, como vimos acima, nos diz que a selecção natural pode tomar um papel forte enquanto fiador epistémico<sup>183</sup>:

«A ideia de que a sobrevivência do mais apto é uma alternativa ao tropismo para a verdade é confusa. Quando a relação entre os dois é apropriadamente formulada torna-se claro que as várias circunstâncias em que a selecção favorece mecanismos cognitivos não-fiáveis todas elas envolvem obter tanta verdade quanto possível dadas as circunstâncias. Todos os processos de selecção são constrangidos, ou de outro modo os organismos “viveriam para sempre, seriam inexpugnáveis aos predadores, depositariam ovos a um ritmo infinito, e por aí adiante” (Maynard Smith 1978, 32). Um sistema cognitivo óptimo não-constrangido teria todas as crenças verdadeiras relevantes para as suas actividades e nenhuma crenças falsas, mas isto não é possível. A evolução seleciona os tropismos

<sup>180</sup> «To suppose that the particular historical narrative adds nothing to a broad ecological model is to suppose that a detailed causal explanation adds nothing to a suggestive idealisation.»

<sup>181</sup> «The classification of a single domain of objects in a series of increasingly abstract vocabularies allows access to generalisations which are hidden at lower levels by the multiple realisability of the kinds that enter into those generalisations.»

<sup>182</sup> «increases in fitness are explained by successful truth-tracking and reductions in fitness by failures in truth-tracking»

<sup>183</sup> No caso de Griffiths & Wilkins, a sua defesa da hipótese A é independente da tese do tropismo autonómico para a verdade, ainda que Griffiths, por sinal, simpatize com os projectos das semânticas de sucesso [em comunicação pessoal].

para a verdade no mesmo sentido que otimiza qualquer traço sob selecção – faz o seu melhor dados os constrangimentos.»<sup>184</sup> [Griffiths & Wilkins 2015:232]

Portanto, a selecção de características que maximizam a fitness não é uma alternativa à selecção de mecanismos cognitivos com tropismo para a verdade, mas como há constrangimentos inerentes aos fenómenos de selecção, é expectável que haja sempre mecanismos cognitivos com tropismos imperfeitos (entenda-se por imperfeitos, com fiabilidade inferior a 100%). Para utilizar uma expressão de Richard Law [1979:379], inspirada pelo ensaio de John Maynard Smith [1978], a *evolução natural* não favorece a emergência de *demónios Darwinianos* (*Darwinian demons*), i.e. de seres capazes de maximizar todos os aspectos da sua fitness simultaneamente e sem constrangimentos [cf. el Mouden *et al.* 2012:29-30]. Demónios Darwinianos maximizariam sem constrangimentos a sua fitness no domínio epistémico. Como já nos dizia Darwin, «como a selecção natural trabalha somente por e para o bem de cada ser, todos os dotes corporais e mentais tenderão a progredir na direcção da perfeição»<sup>185</sup> [Darwin 1859:577], o que significaria que os demónios Darwinianos cognoscentes «teriam todas as crenças verdadeiras relevantes para as suas actividades e nenhuma crença falsa»<sup>186</sup> [Griffiths & Wilkins 2015:232]. Estes demónios não existem por diversas razões. P.ex., a evolução não é apenas determinada pelo mecanismo de selecção natural [Anexo A], e a própria dinâmica da selecção natural apresenta elementos não-adaptacionistas [Anexo B]. Ademais, há ainda constrangimentos à selecção causal que se dão (e.g.) à conta da co-selecção doutras propriedades instanciadas ao mesmo nível de descrição. O tropismo para a verdade é, pois, uma propriedade *ecológica* ou *causa diferencial dos valores de fitness*<sup>187</sup> contra a optimização da qual se dá a optimização de outras propriedades (nenhuma das quais, dizem-nos estes autores, é a falsidade, como veremos de seguida). Dizem-nos então:

«A proposta de que as nossas adaptações cognitivas evoluídas não procuram a verdade não pode significar que elas procuram ao invés a fitness. Deve significar que elas procuram outra propriedade que é uma alternativa genuína à verdade no mesmo nível de explicação. (...) Se elas falharem em procurar a verdade tão eficazmente quanto poderiam é porque têm um tropismo para a verdade sujeito a constrangimentos. A moeda do sucesso evolucionário no domínio da cognição continua a ser um tropismo para a verdade.»<sup>188</sup> [Griffiths & Wilkins 2015:235]

Contudo, e ainda que causa diferencial da fitness, um tropismo para a verdade é ainda uma propriedade bastante abstracta, podendo ser realizada por múltiplos sistemas ou mecanismos

---

<sup>184</sup> «The very idea that fitness tracking is an alternative to truth-tracking is confused. When the relation between the two is properly formulated it becomes clear that the various circumstances in which selection favors unreliable cognitive mechanisms all involve obtaining as much truth as possible given the constraints. All selection processes are constrained, or else organisms “would live for ever, would be impregnable to predators, would lay eggs at an infinite rate, and so on” (Maynard Smith 1978, 32). An unconstrainedly optimal cognitive system would have every true belief relevant to its activities and no false beliefs, but this is not possible. Evolution selects for truth tracking in the same sense that it optimizes any other trait under selection – it does the best it can given the constraints.»

<sup>185</sup> «And as natural selection works solely by and for the good of each being, all corporeal and mental endowments will tend to progress towards perfection.»

<sup>186</sup> «[a]n unconstrainedly optimal cognitive system would have every true belief relevant to its activities and no false beliefs».

<sup>187</sup> O que se coaduna perfeitamente com a nossa tese de que os conteúdos têm propriedades ecológicas [§II.1-2].

<sup>188</sup> «So the proposal that our evolved cognitive adaptations do not track truth cannot mean that they track fitness instead. It must mean that they track some other property which is a genuine alternative to truth at the same level of explanation. (...) If they fail to track truth as effectively as they might, it is because they are tracking truth subject to constraints. The currency of evolutionary success in the domain of cognition remains truth-tracking.»

(e.g. por mecanismos inferenciais dedutivos ou indutivos). Griffiths & Wilkins sugerem que a propriedade de ter um tropismo para a verdade:

«É melhor pensada como uma medida geral de um certo tipo de interacção ecológica com o ambiente, semelhante a “eficiência na procura de alimentos” ou “eficiência respiratória”, como veremos abaixo. A reivindicação de que um organismo tem sucesso porque é melhor que os seus rivais a encontrar a verdade tem tanto conteúdo empírico como a reivindicação de que teve sucesso porque era mais eficiente na procura de alimentos ou tinha um melhor sistema respiratório.»<sup>189</sup> [idem:234]

E isto dá-nos já alguns indícios de propriedades ecológicas que ocorrem ao mesmo nível do tropismo para a verdade e cuja selecção, enquanto tal, constrange a selecção de tropismos para a verdade. P.ex., a optimização do tropismo para a verdade é equilibrada com a optimização paralela do sistema respiratório, e esta por sua vez é também equilibrada com a necessidade de encontrar recursos essenciais à sobrevivência, e por aí adiante. Para além disto, ao custo da optimização do tropismo para a verdade dos nossos mecanismos cognitivos acrescentam-se ainda outro tipo de constrangimentos lógicos, económicos e ecológicos, aos quais dedicaremos o próximo capítulo.

A lição a reter é que não podemos dizer da selecção causal que ela *não* é optimizadora, apenas que a optimização é *constrangida* [el Moulden *et al.* 2012:30]. Portanto, dizer que os sistemas cognitivos têm um tropismo para a verdade optimizado equivale a dizer que os sistemas cognitivos são optimamente fit *dados os constrangimentos* – o que, se a hipótese A for correcta, é sempre verdade enquanto a influência da selecção natural na evolução dos nossos sistemas cognitivos for superior àquela doutros mecanismos [cf. Anexo A], questão empírica cuja resposta determina a verdade das premissas quantitativas 8 e 8\* dos argumentos evolucionários e à qual retornaremos mais adiante como um possível contra-argumento [§V.]. Por enquanto explicitamos unicamente a hipótese A<sup>190</sup>, deixando claras as suas consequências, que melhor nos ajudarão a decidir quanto ao seu valor de verdade.

Uma consequência de interesse é a de que a função cognitiva estará sempre associada a um tropismo para a verdade, independentemente do seu desempenho, e que o desempenho desta função é avaliado, à medida da fitness, pelas suas propriedades ecológicas. Enquanto tal, existirão circunstâncias ambientais nas quais exibir um tropismo para a verdade não é vantajoso, *i.e.* circunstâncias nas quais ter mecanismos cognitivos não é uma boa estratégia ecológica. Isto é tão necessário quanto existirem circunstâncias em que não é vantajoso do ponto de vista evolutivo ser eficiente a respirar por pulmões (e.g. se a atmosfera for anaeróbica) ou na locomoção (e.g. aparentemente as plantas não têm órgãos locomotores proficientes, e alguém com uma deficiência motora poderá executar tão bem ou melhor um “trabalho de secretária” que um corredor olímpico).

Num primeiro instante isto poderá levar-nos a suspeitar que existem circunstâncias nas quais ter um mecanismo cognitivo com um “tropismo para a falsidade” poderia ser mais vantajoso do que ser portador de um mecanismo cognitivo com um tropismo para a verdade. Mas esse raciocínio é erróneo. Para ver porquê bastará pensarmos no caso do desempenho do deficiente motor contra o desempenho do corredor olímpico, tendo em vista a selecção da eficiência

<sup>189</sup> «It is best regarded as a general measure of a certain kind of ecological interaction with the environment, akin to “foraging efficiency” or “respiratory efficiency,” as we will see below. The claim that an organism succeeded because it was better than its rivals at tracking truth has just as much empirical content as the claim that it succeeded because it was a more efficient forager or had a more efficient respiratory system.»

<sup>190</sup> Infelizmente, Griffiths & Wilkins [2015] não discutem as premissas 8 vs. 8\* [cf. §V.].

locomotriz: as circunstâncias em que o deficiente motor será favorecido acima do corredor olímpico são precisamente aquelas em que é completamente irrelevante para a fitness ter um órgão locomotor funcional. Aliás, nessas mesmas circunstâncias, ter um órgão locomotor funcional extremamente eficiente poderá até ser desvantajoso, pois a sua manutenção é custosa, e como a sua utilidade é ausente, o órgão será deletério para a fitness. Como tal, nessas circunstâncias, haverá inclusive selecção causal *negativa* de órgãos locomotores (como a que dá origem a estruturas vestigiais) e não selecção causal *positiva* de um órgão locomotor disfuncional ou não-adaptativo.

O mesmo se passará então com os nossos mecanismos cognitivos em circunstâncias *sui generis*. Não existirão circunstâncias ambientais nas quais há selecção causal positiva de mecanismos cognitivos quando há um tropismo para a verdade, mas, sim, selecção causal *contra* mecanismos cognitivos. Evidência deste tipo de “vestigiação” é difícil de encontrar, dado que até agora os mecanismos cognitivos parece terem-se provado bastante vantajosos face à acrescida complexidade ambiental em todos os taxa animais. Porém, há pelo menos um caso que poderá ser relevante para o ponto que aqui defendemos: os Tunicata ou Urochordata são metazoários que, pertencendo ao Filo *Chordata*, apresentam espinais-medulas; porém, estes só apresentam espinais-medulas na sua fase larvar, durante a qual navegam os mares à procura de uma rocha ou coral onde possam acoplar e sofrer metamorfose para a fase adulta; na metamorfose absorvem o seu gânglio cerebral e espinal-medula que lhes permitia controlar o movimento e que é agora supérfluo (ou representa mesmo um desperdício energético) [cf. Dennett 1991:177]<sup>191</sup>.

Para falsificar a nossa hipótese A terá então de se demonstrar a existência de selecção causal de mecanismos cognitivos com uma propriedade ecológica a que chamaremos tropismo para a falsidade, *viz.* mecanismos cognitivos que, de algum modo, e para além da mera coincidência, são contribuidores sistemáticos para a preservação da autonomia dos biosistemas, apesar de do seu tropismo para a falsidade (não somente como subprodutos ou enquanto fenómeno pleiotrópico) [cf. Anexo B], violando assim a hipótese A, assim como a tese do tropismo autonómico para a verdade [cf. §II.1-2.].

## 5. Constrangimentos Lógicos, Económicos e Ecológicos à Selecção Natural de Mecanismos Cognitivos com Tropismos para a Verdade

Como o título indica, neste capítulo falaremos brevemente de vários tipos de constrangimentos lógicos, económicos e ecológicos à selecção natural de mecanismos cognitivos com tropismos para a verdade, o que nos permitirá colocar de lado alguns contra-exemplos tradicionais apontados aos argumentos evolucionários epistemicamente optimistas, reivindicando a hipótese A.

O nosso ponto de partida, já introduzido [§II.2.], é que todos os sistemas autónomos exploram informação do seu ambiente [Wuketits 1990:53 e Gell-Mann 1994:348], da qual depende a sua sobrevivência em ambientes complexos [Godfrey-Smith 1998 e 2002] e hostis [Sterelny 2003]. Neste sentido, *a epistemologia evolucionária diz respeito à evolução do processamento*

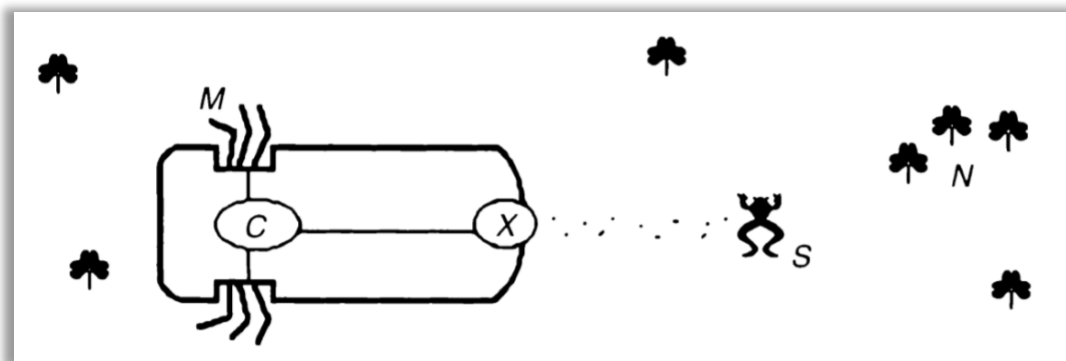
---

<sup>191</sup> Este exemplo também constitui evidência a favor da tese da função mínima da cognição de Jékely *et al.* [2015], cf. §I.4..



de informação [Wuketits *idem*], e os constrangimentos à selecção natural de mecanismos cognitivos com tropismos para a verdade serão melhor entendidos como *constrangimentos ao processamento de informação*.

Introduza-se então um modelo rudimentar e abstracto de um sistema detector de pistas ambientais. O sistema utiliza mecanismos que medeiam as suas interacções com o mundo, detectando pistas do ambiente através de aparatos sensoriais, processando a informação percebida, e executando, ou não, uma acção com base na informação processada. Adoptando o exemplo de Godfrey-Smith [1991:710-1]:



O sistema pode detectar, através do sensor fotossensível *X*, pistas emitidas por *S*, representado uma pista verídica (ou sinal) de um sapo, ou responder igualmente a *N*, que poderá constituir ruído ambiental. Essas pistas são processadas dando origem a *C*, o estado do sistema cognitivo com poderes representacionais (e.g. um estado mental, uma crença), que poderá desencadear um movimento *M*. Se a satisfação da autonomia deste sistema depender de, p.ex., caçar o sapo, então o sistema tem sucesso na sua cognição quando este reage com movimento por resposta a uma pista ambiental emitida por um sapo e não por ruído. A presença do sapo, *S*, e não qualquer outro sinal, *N*, constitui assim as condições ambientais que satisfazem a autonomia de um sistema com uma representação *C*, i.e. *C* é uma representação verdadeira e funcional se e somente se *S* estiver presente. Como nos diz Godfrey-Smith, parece ser o caso que, também de acordo com um modelo de um sistema de detecção de sinais, «a verdade permanece um objectivo claro da cognição»<sup>192</sup> [1991:722; cf. §III.4.], a ser avaliada por padrões correspondentistas em acordo com os quais «acertos ou rejeições correctas são verdades, e falhas ou falsos alarmes são erros» [Godfrey-Smith 1991:722]. Estas quatro interacções possíveis podem ser representadas numa matriz de decisão semelhante à introduzida anteriormente [§III.2.]:

	Sinal (verídico)	Ruído (não-informativo)
Agir	Benefício de uma acção de sucesso, menos o custo da acção.	Sem benefício, mas com o custo da acção. <i>Erros do tipo 1 ou falsos positivos.</i>
Não agir	Sem custo directo da acção, mas com custos variando da perda de oportunidade (a fitness é reduzida relativamente àquela dos organismos que agem) à morte. <i>Erros do tipo 2 ou falsos negativos.</i>	Rejeição correcta, sem benefícios nem custos. (a situação é neutra para a fitness)

<sup>192</sup> «if *C* [the mental state formed on the basis of *X*] is regarded as representing *S*, hits and correct rejections are truths, while misses and false alarms are errors. So truth remains a definite goal of cognition, though the truth-linked virtue of 'reliability' is taken apart and overhauled»

«Custos e benefícios de agir ou não agir em resposta a um sinal verídico ou ruído»  
[Adaptado de Griffiths & Wilkins 2015:240]

Assim, podemos esperar alguma afinidade entre as conclusões base das semânticas de sucesso [cf. §III.2] e a epistemologia evolucionária optimista de Griffiths & Wilkins, que são fundadas numa análise do mesmo tipo de casos<sup>193</sup>. Com base nestas matrizes, podemos também especular que os mecanismos cognitivos que terão maior sucesso serão aqueles que maximizarem o número de detecções de sinais. Um sistema de detecção que produza apenas detecções falsas só poderá aumentar a fitness por acaso, e um sistema que produza sobretudo falsos negativos terá muito provavelmente uma fitness inferior a um sistema que produza sobretudo falsos positivos. De um modo ou de outro, mecanismos cognitivos produtores de falsos positivos e falsos negativos serão sempre acompanhados de um decréscimo na fitness dos seus portadores, pois terão sempre um custo associado à agência em falso, associado à perda de oportunidade, ou, drasticamente, à sua morte. E, à semelhança do que sucede com a estrutura das semânticas de sucesso, «aumentos na fitness são explicados pela detecção com sucesso de sinais e reduções na fitness com falhas na detecção de sinais»<sup>194</sup> [Griffiths & Wilkins 2015:239].

Mas a reintrodução destas tabelas serve-nos outro propósito. Pretendemos elucidar de que modo é que os constrangimentos ao processamento de informação são tais que a optimização dos nossos mecanismos cognitivos nunca os livrará da possibilidade de produção de erros do tipo 1 e 2, o que nos permitirá explicar porque é que, mesmo que a hipótese A seja verdadeira, incorremos ainda assim em tantos erros. Ora, o argumento que aqui apresentamos diz-nos que os constrangimentos se devem às condições ambientais, às limitações do aparato sensorio e do processamento dos estímulos ambientais propriamente ditos, mas não à falsidade da hipótese A.

Um dos motivos pelos quais os sistemas continuarão a produzir representações falsas deve-se ao facto da fiabilidade de muitas das pistas ambientais ser inferior a 1. As representações formadas pelos nossos sistemas cognitivos são muitas vezes o produto de um processamento de informação com um nível de incerteza associado, pelo que as representações em si poderão ser, também elas, mais ou menos fiáveis<sup>195</sup>.

Ademais, a selecção de mecanismos cognitivos epistemicamente fiáveis é *logicamente* constrangida pela assimetria de custos associados aos tipos de erros que podemos cometer. Errar de uma certa maneira (e.g. cometendo erros do tipo 1) poderá ser melhor que errar de outra (e.g. erros do tipo 2), «e é logicamente impossível reduzir a probabilidade de cometer um erro de

---

<sup>193</sup> Relembremos-nos que o princípio de Ramsey é assumido pelas tabelas-matriz de decisão, e que a teoria de verdade por correspondência é compatível com uma semântica de sucesso [cf. §III.4.].

<sup>194</sup> «increases in fitness are explained by successful truth-tracking and reductions in fitness by failures in truth-tracking»

<sup>195</sup> Esta hipótese encontra apoio nas contemporâneas *teorias Bayesianas da mente* ou *da cognição enquanto processamento predictivo*. Estas são abordagens formais ao estudo da cognição de acordo com as quais esta é entendida como processamento predictivo [Friston & Stephan 2007; Friston 2010; e Clark 2015]. Acreditamos que poderá haver uma grande proximidade entre estas teorias e a abordagem interactivista. De facto, algum do trabalho que citamos [e.g. Stephens 2001 e Trimmer & Houston 2014], à semelhança destas teorias, modelam a racionalidade ecológica a partir de estatística Bayesiana. Além do mais, os próprios autores fundadores desta abordagem teórica à cognição partem também do princípio de que os seres vivos auto-organizam-se, com base na teoria termodinâmica (estatística de von Helmholtz).

tipo 1 sem aumentar a probabilidade de cometer um erro do tipo 2 (e vice-versa)»<sup>196</sup> [Griffiths & Wilkins 2015:237-8].

Um exemplo do dia-a-dia permitir-nos-á dar sentido a este constrangimento lógico: os pardais-comuns respondem com fuga a simples movimentos que *poderão* constituir ameaças. Na maior parte das vezes os pardais responderão a movimentos que na realidade não constituem perigo, *i.e.* produzirão falsos positivos. Apercebendo-se desta faceta dos sistemas cognitivos, inúmeros autores congeminaram que, então, na maioria dos casos, os pardais têm representações falsas do mundo, e, o que parece ser pior para o estatuto da hipótese A, os seus mecanismos cognitivos parecem ter sido selecionados por reagirem dessa maneira. Se os pardais não tivessem mecanismos cognitivos assim tão sensíveis às pistas ambientais, se, p.ex., os pardais esperassem mais tempo para confirmar que a pista é emitida por um predador, então não fugiriam com tanta antecipação, o que significaria um aumento da sua susceptibilidade à predação. Donde alguns autores concluíram de uma forma drástica:

«Durante imensos períodos de tempo o intelecto nada produziu para além de erros. Alguns destes provaram-se mais úteis e ajudaram a preservar a espécie: aqueles que os atingiram ou os herdaram tiveram melhor sorte na sua luta por si próprios e pela sua progénie.»<sup>197</sup> [Nietzsche 1974:169/III#110].

Ou, de uma forma menos exacerbada:

«Uma estratégia inferencial muito cautelosa, aversiva ao risco – uma estratégia que salta para a conclusão que o perigo está presente com pouca evidência – irá tipicamente levar a crenças falsas mais vezes, e a verdadeiras menos frequentemente, do que uma menos sensível, que espera por mais evidência antes de emitir um juízo. Não obstante, a estratégia de aversão a riscos não-fiável, propensa a erros, poderá ter sido favorecida pela selecção natural. Pois *a selecção natural não quer saber da verdade*; importa-se apenas com o sucesso reprodutivo. E, do ponto de vista do sucesso reprodutivo, *é melhor estar seguro (e errado) do que arrependido*.»<sup>198</sup> [Stich 1990:62, nossa ênfase].

Por consequência, estes autores, entre outros [e.g. Dennett 1987, Stephens 2001, Sage 2004, McKay & Dennett 2009, e Trimmer & Houston 2014] partilham da crença de que a persistência da selecção de mecanismos cognitivos com *viéses* para a produção de erros de tipo 1 e 2 constitui evidência contra a hipótese A.

Porém, estes *viéses*, que são há muito conhecidos (e.g. a fábula popular *O pastor e o lobo*, de Esopo), não constituem de facto evidência contra a hipótese A. A hipótese de que só há selecção causal positiva de mecanismos cognitivos quando estes têm tropismos para a verdade é compatível com a tese de que devemos «errar pelo lado da prudência»<sup>199</sup> [Dennett 1987:51, n.3]. Do voo do pardal podemos inferir que o seu sistema cognitivo provoca acções quando a

---

<sup>196</sup> «[I]t is logically impossible to reduce the probability of committing a type one error without increasing the probability of committing a type two error (and vice-versa).»

<sup>197</sup> «Over immense periods of time the intellect produced nothing but errors. A few of these proved to be useful and helped to preserve the species: those who hit upon or inherited these had better luck in their struggle for themselves and their progeny.»

<sup>198</sup> «[A] very cautious, risk-averse inferential strategy – one that leaps to the conclusion that danger is present on very slight evidence – will typically lead to false beliefs more often, and true one's less often, than a less hair-trigger one that waits for more evidence before rendering a judgment. Nonetheless, the unreliable, error-prone, risk-averse strategy may well be favoured by natural selection. For natural selection does not care about truth; it cares only about reproductive success. And from the point of view of reproductive success, it is often better to be safe (and wrong) than sorry.»

<sup>199</sup> «to err on the side of prudence»

fiabilidade das pistas detectadas é muito baixa (ou simplesmente incerta) e a utilidade esperada da sua acção é elevada, mas não poderemos inferir que não é epistemicamente fiável. Muitas vezes os sistemas cognitivos causam acções precisamente porque as pistas ambientais não são fiáveis (o que no caso do pardal equivale a dizer que este age sem saber se há verdadeiramente um predador por detrás do arbusto ou não).

Noutro exemplo de Stephen Stich [1990:61-2; retomado por Sage 2004:103], um omnívoro que evita todos os cogumelos sarapintados num ambiente gastronomicamente heterogéneo onde existem mais espécies de cogumelos pode não ser afectado pela prudência da sua decisão. Mas a prudência da sua decisão pode ter consequências drásticas num ambiente homogéneo onde cogumelos pintalgados são dos poucos alimentos abundantes. Como nem todos os cogumelos com pintas são venenosos e há poucas alternativas à alimentação, o ambiente homogéneo favorecerá a evolução de sistemas que processem robustamente a informação antes de emitirem juízos acerca da alimentação, ao invés de favorecer um sistema cognitivo enviesado para a produção de falsos positivos (sendo que um viés para produzir falsos negativos num ambiente como o segundo é mais custoso).

Randolph M. Ness [2005] dá-nos ainda outro exemplo de design de mecanismos de detecção de sinais que vai também ao encontro desta conclusão. Na construção de um detector de fumos evitam-se custos mais elevados se o detector emitir mais falsos positivos do que falsos negativos. Para isso os detectores têm de ser mais sensíveis.

Porém, se pudéssemos desenhar um detector de fumo que estivesse sempre correcto, seria esse que desenvolveríamos. E, de facto, essa é a heurística utilizada para desenvolver protótipos. E se houvesse um mecanismo cognitivo capaz de discernir sempre os cogumelos venenosos dos não venenosos (sem grandes custos associados à detecção), então esse mecanismo seria selecionado. Se os pardais se limitassem a fugir quando há realmente perigo não dispenderiam a energia do voo. Porém, nenhum destes mecanismos está presente devido a inúmeros constrangimentos. A este respeito, William Blake escreveu um provérbio bastante eloquente, no qual nos diz que a «prudência é (...) cortejada pela incapacidade»<sup>200</sup> [Blake 2011:140/1790-3:vs.29]. E até Stich [1990:62] parece ter-se apercebido disto quando escreveu que:

«Poderá ser objectado que na situação que estamos a imaginar, a solução óptima seria um sistema inferencial que obtém sempre a resposta certa. E realmente, se tal opção estivesse disponível e não impusesse grandes custos internos, então poderia bem ser a preferida pela selecção natural»<sup>201</sup>

Assim, nenhum destes exemplos parece contribuir para a falsidade da hipótese A. Mas cremos também que estes tipos de contra-argumentos à hipótese A enfatizam erradamente que os sistemas cognitivos ditos enviesados têm crenças falsas. Isto poderá dever-se à dificuldade, já mencionada anteriormente [e.g. §III.6.(a)], de aferir «quais as inferências doxásticas que podem ser retiradas a partir do comportamento»<sup>202</sup> [McKay & Dennett 2009:501], uma questão sobremaneira mais delicada quando se trata da etologia animal. Retomando o exemplo do pardal com um viés para a produção de falsos positivos, perguntemo-nos o que é que nos garante que o pardal acredita veementemente na proposição (quase sempre) falsa  $p$  “existe um predador por detrás do arbusto em movimento”? Em todo o caso, mesmo que concedamos que o pardal acredita veementemente em  $p$ , tal não nos impediria de avaliar o seu sucesso em acordo com o

---

<sup>200</sup> «Prudence is a rich, ugly old maid courted by incapacity»

<sup>201</sup> «It might be objected that in the situation we are imagining, the optimal solution would be an inferential system that always got the right answer. And indeed, if such an option were available and imposed no great internal costs, it might well be the one preferred by natural selection.»

<sup>202</sup> «The issue here is what doxastic inferences can be drawn from behaviour»

princípio de Ramsey e manter que somente as crenças verdadeiras garantirão *sistematicamente* aumentos na fitness. O valor evolucionário de uma atitude epistémica parece deixar-se definir pelo valor de verdade do conteúdo da proposição acerca da qual versa, ainda que este valor seja desconhecido pelo agente, e este dele permaneça para sempre ignorante.

Os pardais e outros bio sistemas com mecanismos cognitivos têm uma *utilidade esperada* associada às suas acções e agem de acordo com essa utilidade e não somente de acordo com o valor de verdade incerto das suas crenças. Há um ajuste ou balanço entre a fiabilidade das pistas e o valor esperado de uma acção, e há um limiar para além do qual se dá o declínio ou acréscimo da utilidade marginal de qualquer acção, e o seu valor crítico, avaliado pelo próprio agente de acordo com parâmetros variáveis, será aquele na qual a fitness do mecanismo cognitivo é optimizada [cf. Trimmer & Houston 2014:316-7]. Não atendendo aqui às qualidades formais deste processamento, ele tornar-se-á mais óbvio pela consideração doutro exemplo. Ora imagine-se então que Pascal vai a um clube e se vê defrontado com a hipótese de aceitar uma bebida de um estranho. Seja  $P$  a probabilidade de  $s$  “a bebida não está adulterada”. Pascal recebe a oferta e questiona-se, “É  $P$  grande? Se sim, então posso aceitar a bebida. Se não... será  $P$  pequena? Se sim, não aceito esta bebida. Se não... será um falso positivo melhor que um falso negativo? Se sim, aceito a bebida. Se não, não aceito a bebida. É melhor manter-me seguro do que arrepen der-me!”, e a evolução natural concorda.

Enfim, estes constrangimentos de que temos vindo a falar devem-se sobretudo à *lógica* da correcção das nossas crenças, «[m]as a correcção é apenas um dos constrangimentos»<sup>203</sup> [Sober 1981:107], existindo outras limitações à optimização epistémica. Sober [1981], p.ex., sugere que muitos destes constrangimentos são “internos” aos sistemas cognitivos. Ele quer com isto dizer-nos que muitos dos sistemas podem ser expostos aos mesmos *inputs* (informativos) e produzir os mesmos *outputs* (em termos de acção) e ainda assim diferirem substancialmente na sua “*fitness interna*” (a expressão é de Sober). De facto, isto sucede inúmeras vezes (e.g. cefalópodes [cf. Godfrey-Smith 2016] e papagaios [cf. Cantlon & Brannon 2010] têm capacidades aritméticas básicas, mas os seus mecanismos cognitivos diferem uns dos outros substancialmente). Seguindo a sugestão de Stich [1990:60-1], podemos pensar nestas diferenças sob uma metáfora computacional. Há sistemas cognitivos com melhores processadores que outros, com melhor capacidade de memória, que consomem menos energia, etc., mas todos eles capazes de produzir os mesmos resultados [cf. Sage 2004:104 para uma listagem mais extensa de constrangimentos não-lógicos]. Enfim, temos de analisar a fitness dos nossos mecanismos cognitivos para além de uma “caixa preta” com recompensas puramente comportamentais»<sup>204</sup> [Trimmer & Houston 2014:328]. Enquanto tal, e sem por isso desafiar a nossa hipótese A, Stich observa que:

«Se os custos são muito elevados, e se há uma alternativa disponível que faz um menos bom, mas ainda aceitável, trabalho de gerar verdades, então a selecção natural pode preferi-la.» [1990:61].

De facto, este parece ser o caso de todos os mecanismos enviesados e também das heurísticas da nossa racionalidade ecologicamente limitada [cf. Gigerenzer *et al.* 1999]. Assim, e se não é o caso que “a selecção natural não quer saber da verdade”, como nos disse Stich num excerto supra citado (um adágio repetido por Paul Thomson 1995:179 e Sage 2004), então o que é o caso é que, como nos dizem Griffiths & Wilkins [2015], a selecção natural de sistemas com tropismos para a verdade está constrangida: “a selecção natural quer saber, sim, da verdade” [cf. Boudry & Vlerick 2014].

---

<sup>203</sup> «But correctness is only one constraint»

<sup>204</sup> «a “black box” with purely behavioural payoffs»

## 6. Contraexemplos à Ponte Mílvia e Respostas

Neste capítulo avaliaremos a possibilidade de selecção natural de mecanismos cognitivos produtores de crenças falsas, contra a hipótese A.

Uma crença falsa, na esteira de McKay & Dennett [2009:493], é uma crença que «não está correcta em todos os seus particulares. (...) [É] uma crença que em algum grau difere dos factos – isto é, é um estado funcional que endossa um estado particular de coisas que não é obtido»<sup>205</sup>, com a devida ressalva de que quando McKay & Dennett escrevem “estado funcional” não querem com isto dizer que é funcional no mesmo sentido em que antecipações aptas são funcionais [cf. §I.3. e §II.1-2.]. Funcional, aqui, significa apenas “com efeitos causais e operacionais”. Um indivíduo que haja de acordo com uma crença falsa dispõe-se a agir de uma determinada forma, e essa disposição é a “função” a que McKay & Dennett remetem, não a uma função biológica no sentido da teoria etiológica ou interactivista [cf. §I.3.]. De facto, é a própria possibilidade de uma crença falsa desempenhar uma função biológica que está em causa neste capítulo.

Para avaliar essa possibilidade começamos por observar que existem várias formas de um mecanismo cognitivo produzir crenças falsas:

- (a) por acaso;
- (b) devido ao mau funcionamento de sistemas cognitivos;
- (c) naturalmente, porque o sistema cognitivo tem por função biológica produzir crenças falsas.

Imediatamente concluímos que crenças falsas do tipo (b) não nos importarão aqui, pois sistemas cognitivos *disfuncionais* não são causal e positivamente seleccionados, sendo então irrelevantes para a veracidade de qualquer uma das hipóteses quanto ao papel da selecção natural enquanto um fiador epistémico. Relativamente a (a), como «a selecção natural não age sobre particulares anómalos, mas em generalizações fiáveis»<sup>206</sup> [McKay & Dennett 2009:498], crenças falsas do tipo (a) também não nos importam aqui [cf. Stich 1990:123 para uma opinião contrária]. O que procuramos são crenças do tipo (c), pois serão crenças falsas e funcionais que irão falsificar a tese do tropismo autonómico para a verdade. Se um estado de coisas esperado não se verifica, então, diz-nos a tese do tropismo autonómico para a verdade, esperá-lo não é conducente ao sucesso das nossas acções, nem ao sucesso evolucionário. Para podermos falsificar a tese do tropismo autonómico para a verdade, precisamos então de exemplos putativos de crenças falsas que conduzam sistematicamente ao sucesso evolucionário devido às suas propriedades causais<sup>207</sup>.

McKay & Dennett [2009] estabelecem critérios para que uma crença falsa possa ser adaptativa, os quais completamos na seguinte lista:

---

<sup>205</sup> «A misbelief is simply a false belief, or at least a belief that is not correct in all particulars. (...) [A misbelief] is a belief that to some degree departs from actuality – that is, it is a functional state endorsing a particular state of affairs that happens not to obtain.»

<sup>206</sup> «[N]atural selection does not act on anomalous particulars, but rather upon [fitness-]reliable generalizations.»

<sup>207</sup> Propriedades causais essas que serão seleccionadas apesar de não estarem numa relação interna com as propriedades do ambiente, à la tese do tropismo autonómico para a verdade.

1. Deve ser alvo de selecção *causal* positiva (não pode ser, e.g., um subproduto da selecção);
2. Deve *representar erradamente* o estado do Mundo para o seu portador;
3. A *falsidade* da representação é a *propriedade* que a torna *adaptativa*;
4. Tem de ser *sistematicamente* adaptativa.

Quais serão, então, as crenças que satisfazem estes critérios?

(a) *Ilusões positivas*.

McKay & Dennett [2009:494 & 505-7] sugerem que *ilusões* (*delusions*), incluindo “ilusões positivas” e auto-engano constituem «instâncias *bona fide* de crenças falsas adaptativas»<sup>208</sup>, acreditando que elas satisfazem os 4 critérios acima delineados [de Cruz *et al.* 2011:524 concordam]. “Ilusão positiva” é um termo cunhado por Shelley Taylor & Jonathon Brown [1988] para designar «auto-avaliações positivas e irrealistas, percepções exageradas de controlo ou de mestria, e optimismo irrealista»<sup>209</sup> [p.193]. Estas são ditas por estes autores adaptativas e necessárias para a saúde mental, *i.e.* «a mente saudável é auto-enganadora»<sup>210</sup> [Taylor 1989:126].

Aqui deixaremos de lado a questão relativa à necessidade de “ilusões positivas” para a preservação da saúde *mental*, que é ainda largamente discutida<sup>211</sup>. Ao invés, foquemo-nos em dois modos nos quais ilusões positivas podem ser vantajosas para a fitness: «[p]odem levar-nos a empreender acções adaptativas; ou podem mais directamente sustentar ou aumentar a nossa saúde, ou a fitness física no sentido do dia-a-dia» [McKay & Dennett 2009:505]. Fica desde já então claro que “ilusões positivas” podem contribuir positivamente para a fitness de dois modos, um dos quais não se deve às qualidades da crença que a tornam apta a guiar acções, como até agora temos enfatizado, mas às suas qualidades com eficácia sobre a nossa saúde, quaisquer que estas sejam.

A asserção cuja verdade necessitamos de contestar tendo em vista a preservação da veracidade da nossa hipótese A é a de que poderão existir «casos em que indivíduos acreditam falsamente que terão grande sucesso, e nos quais o esforço auto-confiante engendrado por tais crenças falsas leva a um maior sucesso do que aquele que seria obtido caso não tivessem falsamente acreditado nisso»<sup>212</sup> [*Op. cit.*:506].

Fora as críticas já apontadas previamente no contexto de uma semântica de sucesso [cf. §III.6.(g)], temos ainda razões para crer que as “ilusões positivas” não deverão ser interpretadas como crenças falsas. Seguindo a crítica de Owen Flanagan à análise de McKay & Dennett, elas talvez sejam melhor interpretadas como envolvendo expectativas ou esperanças positivas,

<sup>208</sup> «*bona fide* instances of adaptive misbelief»

<sup>209</sup> «unrealistically positive self-evaluations, exaggerated perceptions of control or mastery, and unrealistic optimism»

<sup>210</sup> «the healthy mind is a self-deceptive one.»

<sup>211</sup> Agradeço a David Yates por insistir neste ponto.

<sup>212</sup> «cases where individuals falsely believe that they will attain great success, yet where the confident striving engendered by such misbelief leads to greater success than would have been attained had they not falsely believed.»

«noutras palavras, uma atitude positiva. Esperanças e uma atitude “consigo fazer” não requerem uma crença falsa»<sup>213</sup> [2009:519].

Fazendo jus às críticas de Dweck e de Flanagan com um exemplo proposto por este último [*ibidem*], um lutador de boxe que mantenha a “ilusão positiva” de que talvez consiga ganhar uma luta não está comprometido com a crença plena de que vai ganhar a luta, e, não obstante, essa sua atitude optimista parece fazer a diferença entre *tentar* vencer ou perder a luta. A ideia é que “ilusões positivas” dão conta de casos em que a atitude epistémica tem relevância *causal* para a determinação do estado de coisas do Mundo, uma vez que têm influência *directa* sobre a agência dos sujeitos epistémicos, deste modo contribuindo *indirectamente* para a determinação do valor de verdade das proposições acerca das quais as atitudes versam<sup>214</sup>.

De facto, se já se soubesse à partida o desfecho do combate, *i.e.* se já se soubesse à partida se a “ilusão” é de facto uma ilusão ou não, então não haveria necessidade de combate, não se realizariam torneios, apostas, etc.. De facto, se um dia o lutador encontrar um adversário ridiculamente mais capaz, contra o qual se sabe que não poderá de modo algum ganhar, manter a crença que *vai* vencer ou mesmo que *pode* vencer pode significar uma visita ao Hospital, pelo que seria recomendável ajustar o optimismo ou até a disposição a combater às circunstâncias. Como nos dizem os próprios McKay & Dennett [2009:506] (contra si próprios): «[a]final, não é claro que haja algo adaptativo acerca de tentar e falhar».

A tentativa de avaliar as (des)vantagens de “ilusões positivas” levou vários autores a estudá-las a partir da *teoria da gestão de erros (error management theory)*, que tem por base as dinâmicas elementares de uma teoria da decisão. Esta forma de análise, por consequência, permite-nos abordar qualquer “ilusão positiva” à luz de qualquer outra representação produzida por mecanismos processadores de informação com constrangimentos, chegando a conclusões semelhantes às anteriores [§IV.5.]. Foi assim que Martie Haselton & Daniel Nettle [2006] abordaram as “ilusões positivas”:

«[S]e o custo [evolucionário] de tentar e falhar é baixo relativamente ao potencial benefício [evolucionário] de ter sucesso, então uma crença ilusória positiva não é apenas melhor do que uma crença ilusória negativa, mas também melhor que uma crença não-enviesada»<sup>215</sup> [p.58]

McKay & Dennett criticaram estas conclusões:

«Haselton & Nettle falam de uma crença ilusória positiva como sendo melhor do que uma crença não-enviesada, mas presumidamente o que eles querem dizer é que um sistema de crenças orientado para a formação de crenças ilusórias positivas – assumindo que tais crenças são consistentemente menos prejudiciais para a fitness do que crenças ilusórias negativas – pode ser mais adaptativo do que um sistema de crenças não-enviesado» [2009:506]

---

<sup>213</sup> «Accept that “positive illusions” are states of mind that benefit the consumer, but reject the claim that they (all, most, many) are best interpreted as involving false beliefs, as opposed to having positive expectations and hopes – in other words, a positive attitude. Hopes and a “can do” attitude need not require false belief.»

<sup>214</sup> Repare-se ainda que o sucesso da acção torna-se então dependente da própria atitude epistémica, o que se repercute no valor de verdade do conteúdo das crenças propriamente dito: os casos em que o agente acha que vai ganhar e ganha continuam a ser aqueles em que a sua crença é verdadeira – mantém-se uma relação interna entre a verdade e o sucesso [cf. §III.2.].

<sup>215</sup> «[I]f the [evolutionary] cost of trying and failing is low relative to the potential [evolutionary] benefit of succeeding, then an illusional positive belief is not just better than an illusional negative one, but also better than an unbiased belief.»



Todavia, a crítica de McKay & Dennett incide somente no facto de que a ênfase de Haselton & Nettle deveria recair sobre o *sistema* ou *mecanismo cognitivo* e não sobre as *crenças em si* – o que é verdade –, mas negligencia a hipótese de Griffiths & Wilkins [2015] de que mesmo os mecanismos cognitivos enviesados têm tropismos para a verdade, ainda que constrangidos [cf. §IV.3-5.]. De facto, não é nada claro como é que, tudo o resto sendo igual, um sistema de crenças *não*-enviesado possa ser pior para a fitness do que um sistema que só produz falsos positivos e negativos. Produzir crenças falsas ou ilusões será sempre o subproduto tolerável da selecção causal de mecanismos cognitivos com tropismos para a verdade, podendo dar-se o caso que, sendo o optimismo muitas vezes mais vantajoso do que o pessimismo, de acordo com a assimetria lógica de custos [cf. §IV.5.], os mecanismos com viéses para o optimismo epistémico sejam, regra geral, mais favorecidos.

Porém, se esta abordagem de Haselton & Nettle é correcta e se “ilusões positivas” «surgem de sistemas produtores de crenças falsas com viéses adaptativos, então, crenças falsas individuais acerca do sucesso são, discutivelmente, mais um subproduto tolerável do que uma adaptação» [McKay & Dennett 2009:506], falhando outro dos critérios necessários à falsificação da hipótese A.

Até agora discutimos a possibilidade de “ilusões positivas” serem *guias para acções de sucesso*. Resta ainda averiguar a possibilidade de “ilusões positivas” poderem *determinar directamente a nossa saúde*, «regulando respostas fisiológicas e neuro-endócrinas a circunstâncias de stress»<sup>216</sup> [McKay & Dennett 2009:506].

Um exemplo deste tipo é o do famigerado efeito placebo, que ocorre quando «a mera crença de que a recuperação vem aí pode em si provocar a recuperação»<sup>217</sup> [Humphrey 2004:735], «mesmo se a expectativa for irrealista»<sup>218</sup> [McKay & Dennett 2009:506].

Contrariamente ao muitas vezes sugerido, não é nada óbvio que, de acordo com este modo de influência que uma crença pode exercer na nossa fitness (influenciando a saúde), as propriedades causalmente eficazes da crença sejam o seu valor de verdade [cf. Churchland 2015]. Mais uma vez, parece ser a atitude epistémica de expectativa optimista, neste caso, acerca da recuperação, que é causalmente eficaz. A expectativa de recuperar tem sucesso quando recuperamos, e assim a expectativa de recuperar prova-se verdadeira sempre que recuperamos. Parece-nos então válido tratar a satisfação destas expectativas por analogia com a satisfação de desejos [conforme §III.5.], querendo com isto sugerir que o efeito placebo está mais associado ao *desejo de melhorar* do que à *crença de que vamos melhorar* (sendo o objecto do desejo a saúde).

Há vários motivos para acreditar que o efeito placebo é largamente independente do valor de verdade das nossas crenças e, ao invés, dependente da atitude do agente. P.ex., estudos recentes [Carvalho *et al.* 2016] demonstraram que o efeito ocorre mesmo quando os agentes sabem que estão a receber placebos, ou seja, o efeito ocorre mesmo quando os agentes estão cientes de que o que quer que seja que estão a tomar não constitui um tratamento eficaz para quaisquer que sejam as suas maleitas<sup>219</sup>, o que, assumindo uma racionalidade básica por parte dos agentes, significa que estes nem sempre estão enganados acerca do seu tratamento.

---

<sup>216</sup> «regulating physiological and neuroendocrine responses to stressful circumstances»

<sup>217</sup> «[m]ere belief that recovery is coming can by itself bring the recovery about.»

<sup>218</sup> «[T]he expectation of recovery appears to facilitate recovery itself, even if that expectation is unrealistic. This dynamic may be at work in cases of the ubiquitous *placebo effect*»

<sup>219</sup> Agradeço esta informação a Mariana Pires.

Além disso, e por analogia com a teoria da gestão de erro, é possível pensar no efeito placebo como um fenómeno devido ao normal funcionamento do nosso “sistema de gestão de saúde em situações de incerteza”, uma perspectiva bem articulada por Nicholas Humphrey [2004] (e repetida em McKay & Dennett [2009]). Humphrey começa por verificar que, sendo latente uma auto-cura, o efeito placebo não ocorre imediatamente após ficarmos adoentados. A Psicologia Evolucionária providencia-nos uma hipótese explicativa deste facto: sendo uma resposta imunitária que pode acarretar efeitos secundários, os quais são especialmente indesejáveis em circunstâncias nas quais os indivíduos correm riscos extraordinários, então, na ausência das condições nas quais os indivíduos podem levar a cabo a sua auto-cura mais vale ser prudente e não curar. A título de exemplo, um elemento de uma tribo humana pela época do Pleistoceno beneficiaria de aguardar pelas circunstâncias nas quais estaria salvaguardado de potenciais predadores para então desencadear uma resposta imunitária que envolveria febres altas e inacção. «Quanto melhores as perspectivas de uma recuperação rápida, menos há a ganhar com jogar pelo seguro e continuar doente»<sup>220</sup> [Humphrey 2004:736]. Nas palavras de McKay & Dennett:

«Porque os recursos são limitados, existe valor adaptativo em limitar o gasto de recursos, tal como há valor no gasto em si.

Uma boa gestão económica requer prever o futuro e, como tal, o sistema de gestão de saúde necessitaria de ter em conta qualquer informação disponível relevante para as perspectivas futuras.»<sup>221</sup> [2009:507].

A ideia é que a resposta imunitária é apenas induzida quando há pistas ambientais que asseguram que é seguro desencadeá-la. Os agentes avaliam «a natureza da ameaça, os custos de uma medida defensiva, a expectativa de uma remissão espontânea, evidência de como é que outras pessoas se estão a sair, e a presença de apoio social»<sup>222</sup> [Humphrey 2004:736] de modo a desencadear a resposta imunitária total. E – diz-nos Humphrey inspirado – nada parece constituir melhor apoio do que a medicina moderna e as drogas usualmente prescritas pelos médicos para assegurar o recobro (daí que o efeito placebo seja tão frequentemente associado à crença falsa de que um determinado tratamento é responsável pela recuperação).

Assim, se a hipótese psicológica de Humphrey está correcta, podemos dispensar supostos contraexemplos deste tipo do mesmo modo que dispensámos anteriores:

«O que evoluiu foi um viés para prestar atenção e esperar por sinais de segurança antes de activar uma resposta imunitária total, e estes sinais teriam sido, na sua maioria, verdadeiros prenúncios de segurança (doutro modo o viés não teria sido adaptativo e não teria evoluído). (...) A crença-placebo, como tal, não é uma crença falsa adaptativa – é um subproduto de uma adaptação.»<sup>223</sup> [McKay & Dennett 2009:507]

---

<sup>220</sup> «[T]he brighter the prospects for a rapid recovery, the less to be gained from playing safe and remaining sick.»

<sup>221</sup> «Because resources are limited, there is adaptive value to limiting resource expenditure just as there is value in the expenditure itself. / Sound economic management requires forecasting the future, and thus the health management system would need to take into account any available information relevant to future prospects.»

<sup>222</sup> «nature of the threat, the costs of the defensive measure, the prospects for spontaneous remission, evidence of how other people are faring, the presence of social support»

<sup>223</sup> «What evolved was a bias to attend to and wait for signs of security before triggering a full-bore immune response, and these signs would, in the main, have been true harbingers of security (otherwise the bias would not have been adaptive and would not have evolved). (...) Placebo misbelief, therefore, is not adaptive misbelief – it is a by-product of an adaptation.»

Repetimos então a conclusão anterior de que só há vantagens em evoluir uma tendência a formar atitudes específicas relativamente a conteúdos cujo valor de verdade é incerto se os custos de falhar não forem significativos, sabendo que as vantagens de adoptar essa atitude só se verificam quando os conteúdos acerca dos quais ela versa são também verdadeiros ou quando a expectativa se prova verdadeira.

(b) *O Zoo de Zardini*<sup>224</sup>.

O Zoo de Zardini é um dos contraexemplos fictícios mais interessantes com que nos deparámos. Dizemos fictício porque não parece corresponder a nenhum exemplo que ocorra naturalmente ou artificialmente, ainda que talvez possa ser, em princípio, reproduzível. No Zoo de Zardini realizam-se experiências (eticamente dúbias) de selecção artificial inter-geracional: os monitores do parque reivindicam a sua capacidade de seleccionar causalmente os animais cujos mecanismos cognitivos só produzem crenças falsas, pois enquanto a estes são distribuídas rações, aos restantes animais que aparentam ter crenças verdadeiras de acordo com os desafios em que são testados nada é providenciado (são deixados à fome). Todavia, a alegação dos monitores do Zoo de Zardini encontra dificuldades na sua demonstração.

Parece-nos muito difícil, se não impossível, que os monitores possam testar a validade das crenças dos diversos animais mediante a observação da sua (in)capacidade de resolverem problemas. Isto porque uma situação em que os animais recebem recompensas quando não conseguem resolver os problemas que se lhes colocam é difícil de distinguir de uma situação em que os animais não procuram afincadamente a sua resolução, independentemente do valor de verdade das suas crenças acerca de como resolver os problemas. Se estas duas situações forem de facto relevantemente semelhantes e os monitores alimentarem os animais sempre que eles não resolverem problemas, então eles não estarão de facto a seleccionar mecanismos cognitivos produtores de crenças falsas, mas a seleccionar *contra* mecanismos cognitivos. Isto porque, do ponto de vista evolucionário, o órgão que deixar de desempenhar uma função adaptativa tornar-se-á eventualmente num órgão vestigial (neste sentido, parece também haver alguma semelhança entre as estratégias viáveis do Zoo e a estratégia dos Urochordata previamente mencionados [§IV.4.]).

Contra a hipótese do parágrafo anterior, talvez seja razoável sustentar – ou, para meros efeitos argumentativos assumir – que, para que haja selecção causal negativa de mecanismos cognitivos com tropismos para a verdade, é necessário que os animais do Zoo disponham pelo menos de um sistema cognitivo rudimentar que lhes permita comportarem-se como se tivessem crenças (a hipótese que estamos a assumir é a da tese da complexidade ambiental, de acordo com a qual comportamento complexo requer cognição complexa). Porém, também esta adenda será insuficiente para reclamar a veracidade da reivindicação dos monitores, pois para assegurarem a sua sobrevivência no Zoo de Zardini bastará aos animais fingir que têm crenças falsas – o que é indiscernível a partir do seu comportamento. E se a um primeiro olhar esta hipótese poderá parecer dúbia, a realidade é que há pelo menos uma razão forte para considerá-la. Tendo em conta que o sucesso sistemático no Zoo de Zardini é dependente de não agir como se se soubesse a verdade, então, para que haja selecção causal *sistemática* no Zoo, os mecanismos cognitivos dos animais residentes têm de ser *poderosos o suficiente para identificarem sistematicamente quando é que uma crença é verdadeira ou falsa de acordo com os seus tratadores*, para que possam agir sistematicamente em *desconformidade*. Assim, o Zoo de Zardini poderá estar até a seleccionar mecanismos cognitivos com uma Teoria da Mente interespecífica que lhes permita obter crenças verdadeiras, desde que a selecção destes mecanismos seja acompanhada

---

<sup>224</sup> Esta experiência de pensamento foi-me sugerida por Elia Zardini.

da selecção de um mecanismo volitivo que os iniba de agir de acordo com aquilo que para os tratadores é a resposta apropriada.

Poderemos verificar ainda outra dificuldade que diz respeito ao enquadramento biofísico do Zoo e, em particular, à teoria de funções interactivista, que é a seguinte: se os monitores do Zoo forem interactivistas, eles não poderão dizer dos seus animais que se mantêm vivos que as suas biofunções têm como propósito *não* encontrar as condições do meio nas quais as suas pressuposições preservam a sua autonomia. Afinal, de acordo com a perspectiva interactivista, biosistemas com sucesso sistemático nas suas acções têm necessariamente pressuposições verdadeiras acerca do mundo. Assim, enquanto os animais preservarem a sua autonomia, continuará a haver um *fit* entre as suas pressuposições e o mundo, *i.e.*, as suas pressuposições acerca do mundo continuarão a ser verdadeiras. E as condições do Zoo de Zardini em que os animais permanecem autónomos são aquelas em que os animais agem de acordo com aquilo que os monitores reconhecem como falso, pois, como acabámos de ver no parágrafo anterior, *somente se eles tiverem representações verdadeiras acerca daquilo que os seus tratadores reconhecem como falso de um modo sistemático é que permanecerão vivos*. Assim, repetimos a conclusão anterior de que para haver uma selecção causal positiva e sistemática destes animais, eles têm de ter um tropismo para a verdade, mas optar por agir em “modo contrário a ela” (*i.e.* contrariamente ao modo esperado pelos monitores). O mecanismo responsável pela agência-em-falso seria necessário à sobrevivência no Zoo de Zardini e, curiosamente, *este mecanismo volitivo teria representações verdadeiras* no sentido discutido em §III.5.

Enfim, para dizer que os animais do Zoo de Zardini estão a ser escolhidos por terem crenças falsas, teremos de presumir que temos acesso ao conteúdo bem definido de cada crença animal, pois só assim poderemos determinar o seu valor de verdade. Porém, como isto difere de saber se uma crença é útil ou não *simpliciter* – o que pode ser determinado pelo sucesso de uma acção, assumindo para isso apenas alguns conteúdos básicos para os desejos dos animais (e.g. um desejo inevitável: o animal irá ter fome a dada altura) –, mesmo que alguns animais do Zoo de Zardini tenham sucesso em se alimentar não há muito que possamos dizer acerca da verdade das suas representações. Tendo esta dificuldade em mente, desconfiamos que o Zoo de Zardini tornar-se-ia diabolicamente mais interessante se nele estivessem aprisionados humanos, aos quais podemos perguntar directamente em que é que acreditam. Mas os mais engenhosos habitantes humanos do Zoo de Zardini depressa fingiriam ser estúpidos, não reconhecendo grande utilidade a exercícios sofisticados dos seus mecanismos cognitivos (afinal, é tão mais fácil e gratificante simplesmente falhar nos desafios propostos). Enfim, de acordo com esta interpretação, a sobrevivência dos humanos do Zoo de Zardini não depende do processamento de informação do ambiente, pelo que, a tese da complexidade ambiental estando correcta [cf. §I.4.], pode esperar-se a *involução* da sua sofisticada cognição funcional.

Para concluir, enquanto o Zoo de Zardini prova que pode haver selecção negativa de mecanismos cognitivos quando não há vantagem em mantê-los de um ponto de vista económico, não parece definir condições suficientes para provar que (na situação hipotética que descreve) há selecção positiva de mecanismos cognitivos com tropismos para a falsidade e selecção negativa de mecanismos cognitivos com tropismos para a verdade.



Nas duas últimas secções aplicámos a tese do tropismo autonómico para a verdade à naturalização de uma semântica de sucesso e, após termos extraído dela um corolário com relevância para a teoria evolutiva, defendemos a hipótese de que só há selecção causal positiva de

mecanismos cognitivos quando há tropismos para a verdade, uma vez que não há crenças falsas adaptativas. Concluimos também que, devido à existência de mecanismos cognitivos enviesados, alguns deles produtores de falsos positivos ou negativos, a selecção natural nem sempre será um fiador epistémico robusto (o que é especialmente claro quando falamos de subprodutos da selecção da nossa cognição), o que nos permite em parte explicar porque é que, apesar da tese do tropismo autonómico para a verdade e o seu corolário serem verdadeiras, observamos ainda assim tantos comportamentos que concordam com crenças falsas.

A relação entre uma semântica de sucesso interactivista e a ponte Mílvia é visível quando consideramos as previsões que fazem relativamente à relação entre sucesso interactivo e verdade. A semântica de sucesso é uma abordagem conceptual ao problema da determinação dos conteúdos de representações, e a ponte Mílvia é uma tese empírica que diz respeito ao sucesso evolutivo de representações verdadeiras. Enquanto uma semântica de sucesso explica a relação entre representações verdadeiras e sucesso interactivo dizendo que a partir do sucesso interactivo sistemático é possível inferir o valor de verdade de uma representação, a segunda inverte o *explanans* dizendo-nos que todas as observações realizadas concordam com a hipótese de que somente mecanismos cognitivos com tropismos para a verdade é que têm sucesso interactivo sistemático, ajuizando o que é verdade em acordo com um critério correspondentista. É possível integrar ambas as teorias se, no processo de naturalização de uma semântica de sucesso, admitirmos que, p.ex., uma análise *causal* dos veículos de representação revela os motivos *nomológicos* por detrás da relação *interna* existente entre representações verdadeiras e sucesso interactivo sistemático.

## V. PROLEGÓMENOS A UMA TEORIA NATURALISTA DO SUCESSO DA CIÊNCIA

«Não se pode mais negar que o intelecto humano está peculiarmente adaptado à compreensão das leis e dos factos da natureza, ou pelo menos a alguns deles.»<sup>225</sup>  
[Peirce 1955:213 (CP:2.750)]

Terminamos o presente ensaio com algumas reflexões acerca do contributo das teses aqui defendidas para a formulação de uma explicação do sucesso da Ciência enquanto um fenómeno natural.

Um dos maiores obstáculos à formulação de uma teoria naturalista do sucesso da Ciência que concorde com as teses defendidas neste ensaio está expresso na premissa (8) dos argumentos evolucionários optimistas [cf. §IV.2.]: “a selecção natural é o principal determinante da evolução dos sistemas cognitivos”. Bem poderão os optimistas provar que a selecção natural garante a fiabilidade epistémica, que, na ausência de outras adendas, se a selecção natural não tiver sido o principal promotor da *nossa* evolução cognitiva, nada poderão dizer da fiabilidade epistémica da Ciência que seja informado pela história selectiva<sup>226</sup>. Ademais, podem os mecanismos cognitivos ser fiáveis tanto adaptativamente como epistemicamente no que diz respeito a crenças do senso-comum das quais dependemos no dia-a-dia, mas resta provar que esta fiabilidade se estende ao domínio da produção científica. Estas são questões empíricas acerca das quais decorrem ainda inúmeros debates sobre os quais não temos ainda uma opinião suficientemente bem desenvolvida, motivo pelo qual não propomos aqui uma teoria naturalista do sucesso da Ciência.

Todavia, podemos introduzir o debate propriamente dito, que gira em torno da veracidade da tese *ontológica* a que Steven Orzack & Sober [1994; cf. ainda Sober 1996:50]<sup>227</sup> e Godfrey-Smith [2001b:343-7] chamam de “*adaptacionismo empírico*”:

«Em larga medida, é possível prever e explicar o resultado de processos evolutivos atendendo apenas ao papel desempenhado pela selecção. Nenhum outro factor tem este grau de importância causal.»<sup>228</sup> [Godfrey-Smith 2001b:337]

Existem críticas gerais ao adaptacionismo<sup>229</sup> [e.g. Gould & Lewontin 1979], mas aqui importa-nos especificamente aquelas que têm como alvo o adaptacionismo na cognição. Autores

---

<sup>225</sup> «It can no longer be denied that the human intellect is peculiarly adapted to the comprehension of the laws and facts of nature, or at least some of them»

<sup>226</sup> Se não há uma história selectiva dos nossos mecanismos cognitivos, se eles são, p.ex., somente o produto de mutações neutras para a fitness e de deriva genética, então, mesmo que a selecção natural fosse um fiador epistémico sem constrangimentos (portanto, produtora de demónios Darwinianos epistémicos), nada poderíamos dizer acerca da fiabilidade epistémica dos nossos mecanismos cognitivos. De facto, para a conclusão de qualquer argumento evolucionário acerca da justificação epistémica, nenhuma diferença faria qual das hipóteses, A-D, é verdadeira.

<sup>227</sup> Cf. Brandon & Rausher [1996] para críticas a Sober [1994], e Sober [1996] para respostas a estas críticas.

<sup>228</sup> «To a large degree, it is possible to predict and explain the outcome of evolutionary processes by attending only to the role played by selection. No other evolutionary factor has this degree of causal importance.»

<sup>229</sup> Segundo Lewontin [1997:30], a genética populacional tem enfatizado cada vez mais os efeitos de mecanismos não-selectivos, mas a questão de quanto é que estes mecanismos são responsáveis pela nossa

como Lewontin [1998] fundamentam uma desconfiança generalizada relativamente à possibilidade de formular uma história selectiva da cognição humana. A desconfiança deve-se ao facto de que o comportamento é algo que, historicamente, se torna epistemicamente inacessível (Gould [cf. Pinker *et al.* 1997] expressa ideias semelhantes), e à inexistência de fósseis de sistemas cognitivos, dos quais depende a análise comparativa de homologias filogenéticas, cruciais à determinação de histórias selectivas robustas. Outros autores, como Scott Atran [2005], são mais moderados, acreditando que é mais legítimo falar de histórias selectivas relativamente a alguns domínios da cognição do que outros. Mas dado que nenhuma outra espécie produz Ciência, a maioria aconselha cepticismo relativamente a histórias selectivas da cognição científica. Enquanto parece ser obviamente o caso que (e.g.) as nossas crenças do senso-comum exibem uma correlação forte entre a sua verdade e o sucesso das acções que orientam, não é tão óbvio que o mesmo se possa dizer acerca das nossas crenças estéticas, morais, religiosas, científicas, etc.

A opinião vigente é de que existe um horizonte mais ou menos definido para a fiabilidade epistémica providenciada pela selecção natural. O horizonte epistémico da fiabilidade será então delimitado pelo âmbito de acção da selecção natural enquanto fiador epistémico. Neste contexto, o cepticismo poderá então ser reformulado do seguinte modo: é impossível julgar os mecanismos cognitivos produtores de Ciência como sendo ou não adaptativamente fiáveis, ou, então, simplesmente não houve selecção (causal) natural dos mecanismos cognitivos produtores de crenças científicas, etc.. Thomson reúne muitos destes receios num excerto ilustrativo:

«O facto das nossas variadas capacidades cognitivas terem contribuído para a nossa sobrevivência até agora não é um argumento para que elas continuem a fazê-lo, e, mais importantemente, não é garantia de que estas capacidades tendam a produzir imagens do mundo precisas, ao invés de uma das muitas possíveis imagens do mundo adequadas. Não há nenhum argumento evolucionário que vá da nossa sobrevivência ou ao valor de sobrevivência ou verdade das nossas teorias e não há nenhum argumento que vá do valor de sobrevivência dessas teorias para a sua verdade. Anthony O'Hear resume bem a única conclusão aceitável: "O máximo que pode ser dito de uma perspectiva evolutiva é que ocupamos um nicho ecológico no qual o jogo científico não é imediatamente destrutivo" (1984, p.212).»<sup>230</sup> [Thomson 1995:174]

Acrescentando a isto, Thomson [1995:183] desconfia ainda que a *curta história da Ciência* seja insignificante de um ponto de vista evolucionário, de tal forma que não só histórias selectivas da Ciência não estão justificadas, como também não o estão quaisquer inferências (optimistas ou pessimistas) que se possam fazer relativamente ao seu futuro.

---

evolução permanece em aberto. Mais recentemente, Clatterbuck, Sober & Lewontin [2013] argumentaram que a questão de se a selecção natural domina ou não sobre a deriva genética (ou vice-versa) foi caracterizada como uma falsa questão. Por outro lado, autores como Motoo Kimura [1983] argumentam que ao nível da biologia molecular a evolução é sobretudo levada a cabo por deriva genética e mutações neutras ou, de acordo com Tomoko Ohta [1973], ligeiramente deletérias para a fitness, mas mantém que ao nível fenotípico (no qual discutimos as teses apresentadas neste ensaio) a evolução é sobretudo levada a cabo pela selecção natural [cf. ainda Wade 2005].

<sup>230</sup> « The fact that our various cognitive capacities have contributed to our survival up to now is no argument that they will continue to do so, and more importantly, it is no guarantee that these abilities tend to produce accurate, rather than one of the many possible adequate, pictures of the world. There is no evolutionary argument which goes from our survival either to the survival value or the truth of our theories and no argument from the survival value of those theories to their truth. Anthony O'Hear nicely sums up the only acceptable conclusion: "The most that can be said from an evolutionary point of view is that we occupy an ecological niche in which the scientific game is not immediately destructive" (1984, p.212).»

Mas nós temos motivos para desconfiar desta desconfiança<sup>231</sup>. Primeiramente, da observação de que *qualquer ser humano à face da terra pode fazer Ciência* (se ensinado) podemos derivar a conclusão que quaisquer que sejam os mecanismos cognitivos responsáveis pela produção de Ciência, eles são provavelmente (e pelo menos) tão antigos quanto o último antepassado comum a todos os humanos que presentemente habitam o planeta (naquela que é uma distribuição geográfica dita cosmopolita). Depois, é indiscutível que avanços tecnológicos (permitidos pelo progresso científico) melhoraram substancialmente a condição humana em muitos dos seus aspectos de um modo mais imediato ligados à fitness: os humanos reproduzem-se em maiores números e sobrevivem durante mais tempo à custa de vacinação, antibióticos, motores de combustão, computadores, agricultura<sup>232</sup>, etc.<sup>233</sup>. De uma forma ou de outra, *não é permissível dizer da maioria das tecnologias que elas são neutras para a fitness*. Isto é especialmente claro relativamente àquelas que nos permitem intervir directamente e determinantemente na representação dos genótipos (e.g. tecnologias de edição genética como aquelas baseadas no CRISPR).

Um contra-argumento à fiabilidade adaptativa da Ciência dirá que o “jogo científico” (especialmente a partir da revolução industrial) tem-se tornado bastante desastroso visto de uma perspectiva ecológica: as tecnologias (de que são exemplo emblemático os motores de combustão) são responsáveis pela indução antropogénica de alterações climáticas substanciais (recordamos que existe inclusive um termo informal para mencionar o período geológico que vivemos, o Antropoceno [cf. Steffen *et al.* 2011], e que há quem fale inclusive das alterações climáticas como riscos existenciais [cf. Bostrom 2002]). Mas são também as próprias ciências e tecnologias que permitem que nos apercebamos dos efeitos antrópicos no clima.

Existem várias propostas de histórias selectivas para as origens cognitivas da mente científica, das quais apresentaremos apenas uma chamada “paradigma da Ciência nativa” (*native science paradigm*) [apresentado por Gregory Cajete 2000; cf. ainda Carruthers 2004 e Mithen 2004 para desenvolvimentos desta teoria]. De acordo com Cajete, apesar das diferenças claras entre a cultura científica moderna e as culturas indígenas, a cultura indígena, nativa, ou tribal pré-globalização é caracterizável, à semelhança da prática científica, por experimentação, construção de modelos, de tecnologias, e de explicações dos fenómenos com base na observação.

---

<sup>231</sup> De facto, ambas as teses apresentadas no §II. nos convidam a desconfiar de Thomson. Uma vez que a Ciência é capaz de produzir uma série de verdades aproximadas (e podemos aqui restringir as verdades ao domínio do observável para simplificar), então, em acordo com a Ponte Mílvia [cf. §IV.], ou reconhecemos que a Ciência tem origens adaptativas anteriores à sua produção, ou reconhecemos que é uma exaptação recente (que uma característica não seja uma adaptação em virtude de uma história natural milenar não quer por isso dizer que essa característica não seja hoje eficiente e maximizadora da fitness). Alternativamente, em acordo com a aplicação da tese do tropismo autonómico para a verdade à semântica [cf. §III.], da observação de que (e.g.) a mecânica Newtoniana tem permitido com sucesso ao *H. sapiens* enviar foguetões para o espaço podemos concluir que a mecânica Newtoniana é, pelo menos, aproximadamente verdadeira.

<sup>232</sup> Talvez seja mais correcto considerar a prática de cultivo como uma dimensão tecnológica sobremodo proto-científica até há relativamente poucas centenas de anos.

<sup>233</sup> Alguns autores, reconhecendo este efeito populacional das tecnologias, recordam-nos ainda que nem por isso os cientistas enquanto indivíduos se reproduzem em maiores números que os não-cientistas, mas este argumento tem dois problemas: primeiro, pressupõe que os mecanismos cognitivos dos cientistas diferem substancialmente dos mecanismos cognitivos dos não-cientistas (o que, não sendo necessariamente falso, requer demonstração, dado que parece chocar com a possibilidade de todos os humanos aprenderem Ciência); segundo, pressupõe um conceito de fitness *contra* a persistência de linhagem ao longo do tempo [cf. §I.2.2.] e nega a possibilidade de efeitos emergentes ao nível do grupo, da população regional, ou da espécie, característicos da selecção em múltiplos níveis ou simplesmente da hierarquia dos sistemas autónomos [cf. §I.2.3.] (aliás, se a Ciência é um fenómeno sociocultural, poderá só ser possível em grupos de certas dimensões [cf. Mesoudi *et al.* 2013]).



O trabalho de Louis Liebenberg [e.g. 2013] exemplifica esta história natural da Ciência. Diz-nos ele que os seres humanos caçadores-colectores desenvolveram metodologias apropriadas às suas tarefas enquanto pisteiros (*trackers*) e que dependiam do emprego de estratégias cognitivas familiares à Ciência. Os pisteiros fariam uso daqueles que serão essencialmente os mesmos mecanismos cognitivos por detrás da execução (e.g.) da Física moderna, mecanismos esses fundamentais já para o sucesso do estilo de vida e sociedade do *Homo erectus*. Isto deve-se ao facto de que as pistas deixadas pelas presas podem ser difíceis de encontrar, pelo que os caçadores têm de antecipar com sucesso onde é que estas poderão ser encontradas, para isso necessitando de mecanismos inferenciais tipicamente científicos. P.ex., os pisteiros observam os trilhos e a partir daí formulam hipóteses acerca do movimento do animal com base no conhecimento prévio das suas práticas comportamentais; mas estas hipóteses são por vezes revistas com o aparecimento de novas informações e, estando correctas, levam à descoberta de novas pistas. Para identificar o animal responsável pelo trilho, os caçadores utilizam o método indutivo; para explicar o seu comportamento e fazer novas previsões, os caçadores utilizam o método hipotético-dedutivo, passam por conjecturas e refutações, e empregam ainda o raciocínio abductivo (e.g. quando optam por uma de duas hipóteses relativamente às quais as pistas ambientais não parecem bastar para a escolha).

De facto, muitos caçadores-colectores sobrevivem em ambientes tão distintos como o deserto do Kalahari e o Ártico, mas não o fazem devido às suas condições genéticas adaptadas a esses ambientes *específicos*, mas devido à adaptabilidade da sua condição cognitiva e socio-cultural específica a esses ambientes (também dita *neuroplasticidade*). Quer se chame a este tipo de conhecimento “Ciência” (como o faz Liebenberg) ou “conhecimento popular” (*folk knowledge*) (como o faz Atran), a verdade é que faz uso de muitos dos mesmos mecanismos cognitivos que utilizamos no nosso raciocínio científico.

Hoje em dia existe um acordo crescente acerca de quais são estas estruturas cognitivas cuja história evolutiva explica por que é que todos temos igualmente certas crenças de senso-comum acerca do mundo em que vivemos. P.ex., no que concerne à Biologia popular, parece ser o caso que todos os seres humanos estão naturalmente dispostos a acreditar num certo conceito (intuitivo) de *espécie* [e.g. Atran 1998:548], e esta organização categorial a que poderemos chamar “taxonomia popular” é informativa, permitindo em última instância a realização de inferências fiáveis de um modo sistemático. Considere-se um exemplo de Joseph Henrich [2015]: a partir da observação de que Félix, o gato, gosta de leite, é intuitivo concluir que os gatos gostam de leite e não somente que Félix gosta de leite. Há então traços comuns às histórias adaptativas propostas para todos estes tipos de categorização: propõe-se que estas categorias cognitivas evoluíram na medida em que correspondem largamente a tipos naturais putativos. Parafraseando uma expressão Humeana de Atran, os “hábitos da mente” tendem a ir ao encontro dos “hábitos do mundo”. Esta é uma tese que encontrámos previamente em excertos de Simpson [1963:98] e Quine [1969:126] [cf. §IV.2.] e que foi desenvolvida em maior detalhe por outros autores [e.g. Kornblith 1993 & 2002, Atran 1998, e Pinker 2009].

Esta hipótese é então proposta como uma primeira resposta ao desafio de Kornblith de procurar compreender como é que «criaturas com as faculdades cognitivas que a ciência diz que temos vieram a compreender o tipo de mundo que as ciências geralmente dizem que habitamos» [cf. Preâmbulo]. Ela é também uma resposta possível à dúvida de Max Delbrück [1986:280] relativa à capacidade da mente produzir Ciência apesar de ter evoluído previamente a ela. No fundo, essa dúvida é apenas uma recapitulação daquele que foi chamado (erroneamente) de “paradoxo de [Alfred Russel] Wallace”, que diz respeito à aparente inutilidade evolutiva do intelecto humano [cf. Pinker 2009:300-1], e que surge ainda sob a forma de uma questão mais generalizada numa carta de Darwin [1881] a William Graham, quando Darwin se interroga por que é que deveríamos confiar nas convicções da mente humana, sendo esta descendente da

dos “macacos”. Repetindo a resposta, os nossos antepassados agarraram a oportunidade de recorrer ao pensamento abstracto para aumentar a sua fitness.

Porém, todas as hipóteses selectivas da Ciência têm sido severamente criticadas. A maioria dos autores concordam que a nossa ecologia epistémica é severamente limitada [e.g. Richerson & Boyd 2000:332-3]. Pinker, p.ex., argumenta que o nosso conhecimento popular de Física (*folkphysics knowledge*) não parece coadunar-se bem com a ideia de que «as mentes estão desenhadas para a verdade»<sup>234</sup> [2005:18] (o que leva Pinker [2009:358-9] a rejeitar a hipótese selectiva do paradigma da Ciência nativa). Gerhard Vollmer [1984] diz-nos que o nosso nicho ecológico-cognitivo é povoado por objectos de dimensão média (com uma certa extensão material e temporal) aos quais estamos adaptados. Todavia, a Ciência frequentemente representa objectos de dimensões que quebram as barreiras perceptivas do Mesocosmos: dos quanta a Laniakea, das micro-vidas de muões a processos históricos cosmológicos da ordem dos éones.

É então costume, mesmo entre os mais optimistas dos epistemólogos evolucionários, circunscrever a fiabilidade dos mecanismos cognitivos a condições ambientais *semelhantes* às que em que estes foram seleccionados [e.g. de Cruz & de Smedt 2012:417] (uma ideia hoje em dia central para as teorias da *racionalidade ecológica* de acordo com as quais os mecanismos de decisão dos animais estão adaptados a determinados ambientes e *podem* não funcionar bem fora destes [e.g. Gigerenzer *et al.* 1999; Trimmer & Houston 2014:341]). Como Maynard Smith nos disse, «[s]e um animal for colocado num ambiente que difere grandemente daquele a que está adaptado o equilíbrio é quebrado; um peixe fora de água morre»<sup>235</sup> [1975:15].

Contudo, é necessário determinar o ambiente de referência da fitness. «Pode ser que os procedimentos cognitivos que funcionam bem num domínio também funcionem bem noutro; o *Modus Ponens* pode ser útil para evitar tigres e para fazer física quântica»<sup>236</sup> [Fitelson & Sober 1997:13; cf. ainda Sober 1981 e 1994:30]. Na ausência de uma teoria da relevância causal das características, asserções no sentido de limitar a topologia epistémica permanecerão correlações, *i.e.* insuficientes para concluir a favor ou contra a fiabilidade de um mecanismo cognitivo noutro ambiente epistémico.

---

<sup>234</sup> «The idea that our minds are designed for truth does not sit well with such facts.»

<sup>235</sup> «If an animal is placed in an environment which differs too greatly from that to which it is adapted, the equilibrium breaks down; a fish out of water will die.»

<sup>236</sup> «It may be that the cognitive procedures that work well in one domain also work well in another; Modus Ponens may be useful for avoiding tigers and for doing quantum physics»

Na ausência de uma história selectiva fidedigna, Howard Sankey [2014] e Griffiths & Wilkins [2015] propõem uma ponte *indirecta* que parte do senso-comum para justificar a Ciência<sup>237</sup>. Eles defendem que a Ciência não é «nada mais que senso-comum treinado e organizado»<sup>238</sup> [Huxley 1870:77], «nada mais que o refinamento do pensamento quotidiano»<sup>239</sup> [Einstein 1956], ou, noutras palavras, que «[o] realismo dos cientistas pode ser pensado como uma versão mais sofisticada do realismo de senso-comum»<sup>240</sup> [Giere 2010:4].

Mas o que é o senso-comum? Eles começam por definir senso-comum como crenças quotidianas que guiam a nossa acção mundana [p.213]. Por vários motivos, acontece que a maioria absoluta destas crenças são verdadeiras. Relevantemente, Sankey [2014:14] acrescenta ainda que estas crenças não devem ser confundidas com *crenças amplamente aceites* (*widely held beliefs*), sendo mais próximas daquilo que, diz-nos Sankey, David Armstrong chamava de “*senso-comum de base*” (*bedrock common sense*), Alan Musgrave de “*crenças instintivas*” (*instinctive beliefs*), e o próprio de “*senso-comum básico*” (*basic common sense*) [*ibidem*]. Assim pensadas, crenças de senso-comum têm semelhanças óbvias com *crenças de conhecimento popular*<sup>241</sup>. De facto, Sankey parece sugerir isto mesmo quando nos diz que «o senso-comum (...) é o legado natural da humanidade, e pode ser partilhado com algumas espécies de animais não-humanos»<sup>242</sup> [p.21] – o que remete para a hipótese selectiva.

A reacção inicial mais frequente passa por apontar que a Ciência costuma destronar o senso-comum. Porém, dizem-nos estes autores, isso não parece correcto. A Ciência pode destronar crenças amplamente aceites, sim, mas não elimina as nossas crenças instintivas. Atendendo a um caso de estudo tradicional [Kuhn 1962], Sankey mostra-nos como *a Ciência explica o senso-comum*. Os modelos geocêntricos (de Aristóteles-Ptolomeu) foram substituídos por um modelo heliocêntrico (de Copérnico-Kepler-Galileu) e é costume pensar que o heliocentrismo contradiz o senso-comum. Mas não é nada claro que isto aconteça. Poderá igualmente dizer-se que a nossa experiência de senso-comum permanece a mesma independentemente do modelo (o que equivale à rejeição de uma das versões da tese de que a observação é dependente de teoria [cf. Sankey 2014:19, fn.11]), enquanto a nossa compreensão do que sucede na nossa experiência de senso-comum – essa, sim – é alterada.

Neste sentido, a proposta de Griffiths & Wilkins e de Sankey parece repetir um critério de viabilidade teórica empirista: as teorias que são seleccionadas são aquelas que “preservam os fenómenos”, adequando-se a eles. De facto, Sankey chega mesmo a referir que «[o] mundo do realista acerca do senso-comum é o mundo de coisas vulgares de tamanhos médios»<sup>243</sup>

---

<sup>237</sup> Para ser justo, Griffiths & Wilkins raras as vezes falam de *mecanismos cognitivos*, referindo-se sobretudo à fiabilidade de *crenças* em si. Porém, o que parece ser um problema conceptual com a sua exposição, acaba por não levantar grandes problemas. Independentemente, gostaríamos de sublinhar que não é impossível, mas apenas improvável, que tenhamos uma adaptação para produzir (e.g.) uma teoria específica, tal como não é impossível, mas apenas improvável, que tenhamos uma adaptação para produzir uma crença específica. Exemplos deste tipo são, por exemplo, a fobia a serpentes, *i.e.* a crença de que cobras são perigosas.

<sup>238</sup> «nothing but trained and organized common sense»

<sup>239</sup> «All of science is nothing more than the refinement of everyday thinking.»

<sup>240</sup> «The realism of scientists may be thought of as a more sophisticated version of common-sense realism.»

<sup>241</sup> Por consequência, também a crença de que existem objectos como mentes faria parte do senso-comum, em acordo com a Teoria da Mente da Psicologia popular.

<sup>242</sup> «The common sense (...) is the natural endowment of humankind, and may be shared with some species of non-human animals.»

<sup>243</sup> «The world of the commonsense realist is the world of ordinary middle-sized things»

[2014:21], isto é, é o Mesocosmos de que nos fala Vollmer<sup>244</sup>. Mas são Griffiths & Wilkins quem mais se aproximam de definir o mundo das crenças do senso-comum como o nosso mundo fenoménico. Eles começam por observar que as nossas crenças de senso-comum são expressas por palavras que não são partilhadas por outros animais [pp.213-4], mas depressa se abstraem dessa diferença quando nos dizem que «[a] partir de uma perspectiva evolucionária, o mundo do senso-comum só pode ser visto como o *Umwelt* humano»<sup>245</sup> [p.214], sendo que «o *Umwelt* de uma espécie é o mundo descrito usando as categorias perceptuais e conceptuais disponíveis para aquela espécie»<sup>246</sup> [*idem*] (aquilo que Moreno & Mossió chamaram um “*mundo subjectivo significante*” [cf. §II.1.]).

Se definirmos senso-comum nestes trâmites, então parece ser claro que as crenças de senso-comum devem ser, na sua maioria, pelo menos, aproximadamente verdadeiras. Sankey está bem ciente disto quando apresenta a sua própria versão de um argumento evolucionário optimista:

«O ponto de que o senso comum é vindicado em interação prática com o mundo pode ser estabelecido num contexto naturalista evolutivo. Crenças de senso comum sobrevivem porque têm valor de sobrevivência. Têm valor de sobrevivência porque são, na sua maior parte, verdadeiras. A nossa espécie não poderia ter sobrevivido se a maioria das crenças de senso comum nas quais baseamos a nossa interação quotidiana com o mundo fossem falsas. As crenças falsas não levam sistematicamente a acções de sucesso. Apesar da acção baseada numa crença falsa poder ter sucesso, o risco para a sobrevivência aumenta onde a acção é baseada numa crença falsa. O senso comum tanto promove a sobrevivência como é o resultado de um processo de selecção natural. A nossa sobrevivência constitui evidência da fiabilidade do senso comum.»<sup>247</sup> [Sankey 2014:18]

Este parágrafo, que repete o credo pragmatista [cf. §III.2.], bem poderá ser um protótipo da ponte Mílvia. Mas é a partir daqui que as suas opiniões divergem. De acordo com Sankey, a Ciência explica o senso-comum e, devido à sua proximidade com a empiria, não se lhe contrapõe, pelo que também não o elimina. Griffiths & Wilkins, porém, vão mais longe, dizendo-nos que o senso-comum funda a Ciência, sendo que a Ciência apenas acrescenta conceptualmente<sup>248</sup> àquela que já é a nossa natural predisposição conceptual para o mundo. Os conceitos científicos levam-nos a entender o mundo de um modo distinto, alterando, enquanto tal, as interacções

---

<sup>244</sup> Na realidade, a relação entre senso-comum e o mundo inobservável postulado pela Ciência não é nada óbvia. Creemos que a proposta destes autores tem de ser suplementada com uma teoria epistémica e semântica evolucionária que nos permita fazer sentido dos inobserváveis. Em todo o caso, Sankey não hesita em dizer que o senso-comum vai além do que é observável dado que inclui, por exemplo, a crença de que as coisas continuam a existir quando não as observamos [cf. Clark 1984 e Boulter 2007 para desenvolvimentos desta tese].

<sup>245</sup> «From an evolutionary perspective, the world of commonsense can only be seen as the human *Umwelt*.»

<sup>246</sup> «The *Umwelt* of a species is the world described using the perceptual and conceptual categories available to that species.»

<sup>247</sup> «The point that common sense is vindicated in practical interaction with the world may be set within an evolutionary naturalist context. Commonsense beliefs survive because they have survival value. They have survival value because they are for the most part true. Our species could not have survived if the majority of the commonsense beliefs on which we base our everyday interaction with the world were false. False belief does not systematically lead to successful action. Though action based on false belief may succeed, the risks to survival increase where action is based on false belief. Common sense both promotes survival and is the result of a process of natural selection. Our survival constitutes evidence of the reliability of common sense.»

<sup>248</sup> Nós arriscamos a dizer que acrescenta também *perceptualmente* na medida em que nos providencia novos instrumentos capazes de novas detecções, mas eles não fazem menção a esta possibilidade.

possíveis que temos com ele, isto é, o que a Ciência faz é alterar a nossa experiência do *Umwelt* (que, relembramo-nos, eles equacionam com o senso-comum).

A hipótese de Griffiths & Wilkins é de que os nossos esquemas conceptuais ainda estão a evoluir [p.217] e que, porque temos um tropismo para a verdade no que concerne os mecanismos cognitivos produtores de crenças do senso-comum, também conseguimos derivar esquemas conceptuais fiáveis que a ele se aplicam<sup>249</sup>. A Ciência tem de «enfrentar o mesmo tipo de escrutínio do senso comum que qualquer outra adição às nossas crenças»<sup>250</sup> [p.219]. Porque eles argumentem neste sentido e simultaneamente defendem que as faculdades cognitivas dos humanos (e.g.) «não foram selecionadas pela sua capacidade para o cálculo»<sup>251</sup> [p.218], não podemos estender a ponte Mílvia *directamente* à Matemática. *Mutatis mutandis* para a Ciência em geral. Enquanto Sankey parece mais disponível a aceitar a existência de uma história selectiva para as faculdades cognitivas produtoras de Ciência, Griffiths & Wilkins dizem-nos que a Ciência é apta, mas não adaptada.

Uma das razões que nos dão para suspeitarmos da Ciência enquanto fruto de adaptações é a mesma dos anti-adaptacionistas: à falta de homologias não podemos proceder a uma análise comparativa. Isto fica claro quando eles nos dizem que «[n]ão há precedentes noutros animais para mudanças radicais na estrutura do *Umwelt* causadas por inovações cognitivas individuais que se espalham por difusão cultural»<sup>252</sup> [Griffiths & Wilkins 2015:218]. Outra razão é dada pela observação de que incorremos em imensas falácias quando utilizamos – para continuar no mesmo exemplo – cálculo (e.g.) probabilístico (e.g. a falácia da frequência de base). Porém, cremos que este raciocínio é erróneo. Primeiro, repetimos que todos os seres humanos parecem ser capazes de aprender cálculo. Por outro lado, a existência de erros do tipo das que nos falam é, em acordo com a sua própria teoria [cf. §IV.3-5.], compatível com a existência de uma história selectiva de mecanismos cognitivos produtores de Ciência com um tropismo para a verdade: falácias probabilísticas poderão talvez explicar-se em virtude de enviesamentos justificados por constrangimentos ecológicos. Assim, poderemos equiparar a emergência do cálculo a um aumento da expressividade (fenotípica) dos mesmos mecanismos cognitivos com uma história selectiva própria, talvez proporcionada pelo efeito combinado da evolução biológica e *sociocultural*<sup>253</sup>. Mas Griffiths & Wilkins não parecem concordar com esta possibilidade, pelo que lhes resta ainda assim explicar como é que a Ciência pode ser justificada pelo senso-comum enquanto continuarmos a incorrer em falácias.

A solução de Griffiths & Wilkins [2015:219] vai ao encontro das propostas anteriores de Helen de Cruz *et al.* [2011] e de Cruz & Johan Smedt [2012]. Estes autores acreditam que a transmissão cultural de conhecimento científico pode ser enviesada e ainda assim ter um tropismo para a verdade, quiçá, independente do tropismo biológico propriamente dito, e de tal forma que um mecanismo cognitivo enviesado pode participar num processo de selecção cultural que

---

<sup>249</sup> Wimsatt [2007b:209-11] oferece uma análise da compatibilidade destas linguagens enquanto descritivas de níveis diferentes de abstracção que, enquanto tal, acreditamos ir ao encontro da posição de Griffiths & Wilkins.

<sup>250</sup> «[the data, arguments, and methods] have to be able to stand up to the same kind of commonsense scrutiny as any other addition to our beliefs.»

<sup>251</sup> «they were not selected for their ability to do calculus»

<sup>252</sup> «There is no precedent in other animals for radical changes in the structure of the *Umwelt* caused by individual cognitive innovations that spread by cultural diffusion.»

<sup>253</sup> Todavia, uma epistemologia evolucionária tem de dar conta do ser humano não só nas suas dimensões biológicas, mas *socioculturais* [Campbell 1974:413] e, apesar deste truísmo, não abordámos as teorias de evolução sociocultural que se encontram presentemente em desenvolvimento [Mesoudi 2016a, 2016b, & 2017], mas é certo que elas afectam tanto a dimensão semântica como a dimensão epistémica da nossa tese.

corrige os viéses. Aliás, essa será, discutivelmente, uma das próprias funções a que a Ciência se dedica com a ajuda adicional de uma divisão cognitiva de trabalho que lhe é tão característica [Kitcher 1990]<sup>254</sup>. Como nos dizem de Cruz *et alia*:

«O facto de os psicólogos cognitivos (e.g. Kahneman et al. 1982) conseguirem reconhecer as fragilidades do raciocínio humano, e de sujeitos em experiências psicológicas terem melhores desempenhos depois dos seus erros lhes terem sido apontados, sugere que viéses e heurísticas não são tão prevalentes que enevoem por completo o nosso raciocínio, e que o raciocínio defeituoso é frequentemente corrigível com algum esforço mental.»<sup>255</sup> [2011:525]

Porém, enquanto Griffiths & Wilkins e Vlerick & Broadbent [2015] julgam que a auto-avaliação deve ser tão fiável como o senso-comum, para nós não é claro que a capacidade de reflectir acerca das nossas próprias capacidades seja concomitante com a capacidade de as exercer. Trata-se da diferença entre uma capacidade cognitiva e outra metacognitiva [como abordada previamente, §II.1.].

Enfim, apesar do seu tropismo para a verdade, pode dar-se o caso dos mecanismos cognitivos produzirem múltiplas crenças falsas. *Assim, nem todos os mecanismos cognitivos com um tropismo para a verdade serão garantidamente epistemicamente fiáveis*. Primeiramente assegurada a adaptatividade da Ciência, podemos depois pesar até que ponto os nossos mecanismos cognitivos são epistemicamente fiáveis. Porquanto a metodologia científica permite corrigir viéses, diminuindo o nível de incerteza associado aos sinais detectados, e a tecnologia permite a computação morosa dos sinais, a Ciência poderá estar munida precisamente daquilo que fará dos próprios mecanismos cognitivos produtores de Ciência (potencialmente) adaptativamente fiáveis também epistemicamente fiáveis [cf. Griffiths & Wilkins 2015:213].



---

<sup>254</sup> Imaginando que a selecção natural não foi nem é um mecanismo muito presente na evolução humana, poderemos ainda assim averiguar se os mecanismos da evolução cultural contribuem ou não, em média, positivamente para a fitness. Assim, se a selecção natural enquanto fiador epistémico fiável não é a grande responsável pela nossa actual disposição cognitiva, poderemos ainda assim estar justificados em acreditar nos seus produtos devido a um outro fiador epistémico natural – a selecção cultural. Esta é uma hipótese de investigação para uma *epistemologia evolucionária das teorias científicas* [sensu Bradie 1986], mas neste ensaio estamos a trabalhar uma hipótese de *epistemologia evolucionária dos mecanismos cognitivos*.

<sup>255</sup> «The very fact that cognitive psychologists (e.g., Kahneman et al. 1982) are able to recognize the frailties of human reasoning, and that subjects in psychological experiments perform better when their errors have been pointed out to them, suggests that biases and heuristics are not so pervasive as to cloud our reasoning completely, and that faulty reasoning is often corrigible with some mental effort.»

## Anexo A – Como a Evolução Previne o Surgimento de Demónios Darwinianos: Mecanismos Não-Selectivos

Esta é uma lista de mecanismos evolutivos para além da selecção natural.

### (a) *Mutagenese.*

Apesar de existirem diferentes mecanismos responsáveis pela variação da frequência de alelos, a mutação, induzida ou não, aleatória ou não, é a única fonte de novas mutações, sendo que a mutagenese ocorre sem a interferência da selecção natural. Como a maioria das mutações ocorre aleatoriamente (com algumas excepções de que é exemplo a mutagenese artificialmente induzida pela engenharia genética), não há razão para esperar uma correlação entre novas mutações e aumentos na fitness. Pelo contrário, a maior parte dos autores assume que a maior parte das mutações são deletérias ou neutras para a fitness quando não induzidas.

### (b) *Deriva genética.*

A deriva genética é uma força que age sobre as frequências genotípicas em populações finitas independentemente dos seus efeitos fenotípicos. Usualmente ocorre em virtude de amostragens aleatórias e os seus efeitos dependem da representação inicial de cada alelo na população. Quando um alelo tem pouca representação a magnitude da deriva será superior ao longo das gerações, e, por contraste, quando um alelo tem muita representação o efeito será menos pronunciado, de acordo com a lei dos grandes números. De acordo com a lei dos grandes números, com o aumento do número de experiências, mais próximo será o resultado do valor esperado. Diferenças em probabilidade tornam-se mais e mais evidentes com o aumento da população e do número de experiências. Aplicando o teorema ao estudo da frequência de genótipos chega-se a um slogan simples: o tamanho da população importa. Os efeitos da deriva são, enquanto tal, uma questão probabilística, pelo que não é permissível fazer previsões quanto à deriva aumentar ou diminuir a fitness dos indivíduos – pode contribuir igualmente para ambos os efeitos. De facto, uma característica mais fit pode estar pouco representada numa população pequena inicial e ao longo das gerações acabar por desaparecer devido à deriva genética, mesmo sendo a característica mais fit. Como as populações não são infinitamente grandes, a deriva genética tem sempre efeitos [cf. Sober 1996:43].

### (c) *Fluxo génico.*

O fluxo génico ocorre quando existe transferência de genes de uma população para outra, alterando a frequência dos genes em ambas as populações, por vezes introduzindo novas variantes no fundo genético (razão pela qual alguns autores consideram-no como mutagenese). Ocorre principalmente entre populações de uma mesma espécie (e.g. quando pólen ou uma semente é levado pelo vento através dos mares), mas também pode atravessar as barreiras das espécies (e.g. através da transferência lateral ou horizontal de genes), dando origem a evolução reticulada, bastante crucial na evolução de organismos unicelulares (e.g. é usualmente apontado como o fenómeno responsável pela propagação de resistência antibiótica). Apesar de alguns casos como o do exemplo anterior demonstrarem a possibilidade do fluxo génico conduzir ao desenvolvimento de adaptações, nem sempre é esse o caso. Além do mais, alterações nas frequências de alelos resultantes de fluxo génico são também sensíveis à deriva genética e à selecção natural à semelhança de qualquer outra mutação.

Existe ainda a possibilidade de algumas adaptações se seguirem naturalmente das dinâmicas de auto-organização inerentes aos biosistemas, não dizendo respeito a nenhum dos mecanismos (a-c). Discute-se também ainda se a evolução sociocultural merece um lugar nesta lista [cf. Laland & Brown 2011].

## Anexo B – Como a Evolução Previne o Surgimento de Demónios Darwinianos: Selecção Não-Adaptacionista

Esta é uma lista breve de alguns exemplos de fenómenos selectivos que são parcialmente não-adaptativos.

### (a) *Superioridade ou vantagem heterozigótica e outros arranjos.*

Imagine-se um simples cenário Mendeliano no qual genótipos heterozigóticos têm maiores valores de fitness do que homozigóticos dominantes ou recessivos. Neste cenário, para que haja selecção dos fenótipos codificados pelo heterozigótico *Aa* terá de haver selecção simultânea de uma certa frequência de *AA* e de *aa* cuja fitness é inferior. Este tipo de selecção, comumente referida como selecção de equilíbrio, é concomitante com polimorfismos genéticos.

### (b) *Pleiotropia.*

Fenómenos pleiotrópicos dão-se sempre que um gene ou arranjo de genes codifica diferentes respostas fenotípicas, *i.e.* quando a relação entre genótipo-fenótipo é de um-para-muitos [a expressão é de Thomson 1995:171], e estes diferentes fenótipos diferem relativamente ao seu valor de fitness. Fenómenos pleiotrópicos podem diferir no seu valor de fitness, mas ainda assim ser adaptativos em respeito a todos os seus alelos, mas os casos mais interessantes são aqueles em que pelo menos um dos alelos é deletério ou neutro para fitness e outro que co-ocorre com ele contribui significativamente para o aumento da fitness. Nestes casos, à semelhança do caso anterior (a), não pode haver selecção de um dos alelos sem que haja selecção simultânea do(s) não-apto(s).

### (c) *Subprodutos.*

Algumas das características dos organismos podem ser como são por motivos independentes dos seus valores de fitness. Algumas destas características são subprodutos da selecção de outras características. Gould & Lewontin [1979] apelidaram estas características de *enjuntas* ou *penachos* (*spandrels*), fazendo uso de uma metáfora arquitectónica em alusão à Catedral de São Marcos e as suas juntas ornamentadas. As juntas em si não têm nenhuma função própria, recebendo a sua forma tipicamente triangular em virtude dos arcos e tectos que usualmente conectam. Algumas características anatómico-somáticas são também assim devido a constrangimentos ao seu desenvolvimento e nada mais.

Existem outros exemplos deste tipo (e.g. boleia genética (*genetic hitchhiking*), ligação genética (*genetic linkage*), deriva meiótica (*meiotic drive*), e provavelmente alguns que presenteiramente desconhecemos.



## Referências Bibliográficas <sup>256</sup>

- ABRAMS, Marshal; [2009] «What Determines Biological Fitness? The Problem of the Reference Environment» (pp. 21–40), *Synthese*, 1 (166).
- ALDER, Mike; [2004] «Newton's Flaming Laser Sword, Or: Why Mathematicians and Scientists don't like Philosophy but do it anyway» (pp.29–33), *Philosophy Now*, 46.
- ALLEN, Colin; [2003] «Teleological Notions in Biology», *The Stanford Encyclopedia of Philosophy*, Winter 2009 Edition, Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2009/entries/teleology-biology/>>.
- AMUNDSON, Ron & LAUDER, George V.; [1994] «Function without purpose» (pp. 443–469), *Biology and Philosophy*, 4 (9).
- ANDREAS, Holger; [2013] «Theoretical Terms in Science», *The Stanford Encyclopedia of Philosophy*, Fall 2017 Edition, Edward N. Zalta (ed.), forthcoming URL = <<https://plato.stanford.edu/archives/fall2017/entries/theoretical-terms-science/>>.
- ARDILA, Alfredo; [2016] «The Evolutionary Concept of “Preadaptation” Applied to Cognitive Neurosciences» (pp.1-3), *Frontiers in Neuroscience*, 10 (103).
- ARIEW, André & LEWONTIN, Richard C.; [2004] «The Confusions of Fitness» (pp.347-63), *British Journal of Philosophy of Science*, 55.
- ASHBY, W. Ross; [1947] «Principles of the Self-Organizing Dynamic System» (pp.125–128), *Journal of General Psychology*, 37.
- ATRAN, Scott; [1998] «Folk Biology and the Anthropology of Science: Cognitive Universals and Cultural Particulars» (pp.547-609), *Behavioral and Brain Sciences*, 21.
- \_\_\_\_; [2005] «Adaptationism for Human Cognition - Strong, Spurious or Weak» (pp. 39–67), *Mind & Language*, 20 (1).
- BARBIERI, Marcello; [2008] «Biosemiotics: a new understanding of life» (pp.577-99), *Naturwissenschaften*, 95 (7).
- BARRETT, Deirdre; [2010] *Supernormal Stimuli: How Primal Urges Overran Their Evolutionary Purpose*, W.W. Norton and Company, New York, NY.
- BERMÚDEZ, José Luis; [2003] *Thought Without Words*, Oxford University Press, Oxford.
- BERMÚDEZ, José Luis & CAHEN, Arnon; [2015] «Nonconceptual Mental Content», *The Stanford Encyclopedia of Philosophy*, Fall 2015 Edition, Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2015/entries/content-nonconceptual/>>.
- BICKHARD, Mark H.; [2001] «Function, anticipation, representation» (pp. 459-469), *AIP Conference Proceedings*, 573.
- \_\_\_\_; [2003] «Some notes on internal and external relations and representation» (pp. 101-110), *Consciousness and Emotion*, 4 (1).
- \_\_\_\_; [2009] «The interactivist model» (pp.547-591), *Synthese* 166.
- BLACK, Max [1954] «Inductive Support of Inductive Rules» (pp. 191-208), *Problems of Analysis*, Cornell University Press, Ithaca.
- BLACKBURN, Simon; [2005] «Success Semantics», Hallvard Lillehammer & D. H. Mellor (eds.) *Ramsey's Legacy*, Oxford University Press, Oxford.
- BLAKE, William; [2011 (ca. 1790-3)] «Proverbs of Hell», in *The Marriage of Heaven and Hell*, in Alicia Ostriker (ed.), *The Complete Poetry of William Blake*.

---

<sup>256</sup> Todos os artigos referenciados foram consultados por inteiro. Nem todos os livros referenciados foram consultados por inteiro, mas, sim, por vezes, com respeito a capítulos específicos.

BONCOMPAGNI, Anna; [unpublished manuscript from a 2016 talk] «What's the Use of Usefulness? From Pragmatism to Success Semantics and Return», «The Future of Pragmatism», CSMN Workshop 17-18 February 2016, University of Oslo.

BOSTROM, Nick; [2002] «Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards» (pp.31-3), *Journal of Evolution and Technology*, 9 (1).

BOUCHARD, Frédéric; [2008] «Causal Processes, Fitness, and the Differential Persistence of Lineages» (pp.560-570), *Philosophy of Science*, 75 (5), Cristina Bicchieri & Jason Alexander (eds.), *Proceedings of the 2006 Biennial Meeting of the Philosophy of Science Association, Part II: Symposia Papers*.  
\_\_\_\_; [2011] «Darwinism without populations: a more inclusive understanding of the "Survival of the Fittest"» (pp.106–114), *Studies in History and Philosophy of Biological and Biomedical Sciences*, 42.

BOUDRY, Maarten & VLERICK, Michael; [2014] «Natural selection does care about truth» (pp. 65-77), *International Studies in the Philosophy of Science*, 28 (1).

BOULTER, Stephen; [2007] «The "Evolutionary Argument" and the metaphilosophy of commonsense» (pp.369–382), *Biology and Philosophy*, 22.

BOURRAT, Pierrick; [2014] «From survivors to replicators: evolution by natural selection revisited» (pp.517–38), *Biology & Philosophy*, 29.

BRADIE, Michael; [1986] «Assessing evolutionary epistemology» (pp.401-459), *Biology and Philosophy*, 1 (4).

BRANDON, Robert; [1994] «Unsuccessful Semantics» (pp. 175-178), *Analysis*, 54 (3).

BRANDON, Robert; [2006] «Teleology in self-organizing systems» (pp. 267-281), B. Feltz, M. Crommelinck, & P. Goujon (eds.), *Self-Organization and Emergence in Life Sciences*, Dordrecht, Springer.  
\_\_\_\_; [2008] «Natural selection», *The Stanford Encyclopedia of Philosophy*, Spring 2014 Edition, Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/spr2014/entries/natural-selection/>>.

BRANDON, Robert & RAUSHER, Mark; [1996] «Testing adaptationism – A commentary on Orzack and Sober» (pp. 189-201), *The American Naturalist*, 148 (1).

BROOKS, Rodney A.; [1991] «Intelligence without Representation» (pp. 139–159), *Artificial Intelligence*, 47.

CAJETE, Gregory; [2000] *Native Science: Natural Laws of Interdependence*, Clear Light Publishers.

CAMPBELL, Donald T.; [1974] «Evolutionary Epistemology» (pp. 413-63), P. A. Schilpp (ed.), *The Philosophy of Karl Popper*, Open Court, LaSalle, IL. E Gerard Radnitzky, William Warren Bartley, & Karl Raimund Popper (eds.), *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge*.

CAMPBELL, Donald T. & BICKHARD, Mark H.; [2003] «Variations in Variation and Selection: The Ubiquity of the Variation-and-Selective Retention Ratchet in Emergent Organizational Complexity» (pp. 215-282), *Foundations of Science*, 8 (3).

CANTLON, Jessica F. & BRANNON, E.M.; [2010] «Animal Arithmetics» (pp. 55-62), Clayton, N. (ed.), *Encyclopedia of Animal Behavior*, Elsevier Press, Oxford.

CARRUTHERS, Peter; [2004 (2002)] «The roots of scientific reasoning: infancy, modularity and the art of tracking» (pp.73-95), Peter Carruthers, Stephen Stich, & Michael Siegal (eds.), *The Cognitive Basis of Science*, Cambridge University Press, Cambridge.

CARVALHO, Cláudia et al.; [2016] «Open-label placebo treatment in chronic low back pain: a randomized controlled trial» (pp. 2766-2772), *PAIN*, 157 (12).

CHAKRAVARTTY, Anjan; [2007] *A Metaphysics for Scientific Realism – Knowing the Unobservable*, Cambridge University Press, Cambridge.

CHURCHLAND, Patricia; [2015] «Brain Chemicals Explain the Power of Placebos - How chemicals made in our brains reduce pain», *The Wall Street Journal*, 5/August.

CHURCHLAND, Paul; [1985] «The ontological status of observables: in praise of superempirical virtues» (pp.35-47), Paul Churchland & Clifford A. Hooker (eds.) [1985].

- CHURCHLAND, Paul M. & HOOKER, Clifford A. (eds.); [1985] *Images of Science: Essays on Realism and Empiricism, with a Reply from Bas C. van Fraassen*, The University of Chicago Press, Chicago.
- CHRISTENSEN, Wayne D.; [1996] «A complex systems theory of teleology» (pp.301-320), *Biology and Philosophy*, 11 (3).
- \_\_\_\_; [2010] «The Decoupled Representation Theory of the Evolution of Cognition—A Critical Assessment» (pp.361-405), *The British Journal for Philosophy of Science*, 61 (2).
- CHRISTENSEN, Wayne D. & BICKHARD, Mark H.; [2002] «The Process Dynamics of Normative Function» (pp.3-28), *Monist*, 85 (1).
- CHRISTENSEN, Wayne D. & HOOKER, Clifford A.; [2000] «Anticipation in Autonomous Systems: Foundations for a Theory of Embodied Agents» (pp.135-154), *International Journal of Computing Anticipatory Systems*, 5.
- CLARK, Andy; [1984] «Evolutionary Epistemology and Ontological Realism» (pp.482-490), *The Philosophical Quarterly*, 34 (137).
- \_\_\_\_; [2015] *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*, Oxford University Press, Oxford.
- CLATTERBUCK, Hayley, SOBER, Elliott & LEWONTIN, Richard; [2013] «Selection never dominates drift (nor vice versa)» (pp. 577–592), *Biology & Philosophy*, 28.
- COLLIER, John; [2003] «Interactively open autonomy unifies two approaches to function» (pp. p.228-235), Daniel M. Dubois (ed.), *Proceedings of the Computing Anticipatory Systems: CASYS'03 - Sixth International Conference, AIP Conference Proceedings*, 718. New York: American Institute of Physics, 2004.
- CORDERO, Alberto; [2013] «Naturalism and Scientific Realism» (pp.1-14), J.I Galparsoro & Alberto Cordero (eds.), *Reflections on Naturalism*, Sense Publishers, Boston.
- \_\_\_\_; [2015] «On Scientific Realism and Naturalism» (pp.31-43), *Journal of Philosophical Research*, 40.
- \_\_\_\_; [2017] «Retention, Truth-Content and Selective Realism» (pp.245-256), Evandro Agazzi (ed.), *Varieties of Scientific Realism: The Problem of Objectivity and Truth in Science*, Springer.
- CORNING, Peter A. & SZATHMÁRY, Eörs; [2015] «"Synergistic selection": A Darwinian frame for the evolution of complexity» (pp. 45–58), *Journal of Theoretical Biology*, 371.
- COUCH, Mark B.; [2011] «Causal Role Theories of Functional Explanation», James Fieser & Bradley Dowden (eds.), *Internet Encyclopedia of Philosophy*, URL = <<http://www.iep.utm.edu/func-exp/>>.
- DAMÁSIO, António; [2010] *Self Comes to Mind: Constructing the Conscious Brain*, Vintage Books.
- DARDEN, Lindley & MAULL, Nancy; [1977] «Interfield Theories» (pp.43-64), in *Philosophy of Science*, 44 (1).
- DARWIN, Charles; [1858] «To Charles Lyell», 18th June 1858, *Darwin Correspondence Project*, "Letter no. 2285", <http://www.darwinproject.ac.uk/DCP-LETT-2285> [consultado pela última vez a 14 de Novembro de 2017].
- \_\_\_\_; [1961 (1859)] *On the Origin of Species by Means of Natural Selection*, reprinted, Collier, New York.
- \_\_\_\_; [várias] Online Variorum of Darwin's *Origin of Species*, in <http://darwin-online.org.uk/Variorum/>, edi. Barbara Bordalejo (last consulted on the 2<sup>nd</sup> of May 2017).
- \_\_\_\_; [1871] *The Descent of Man, and Selection in Relation to Sex*, John Murray, London.
- \_\_\_\_; [1881a] «To William Graham», 3rd July 1881, *Darwin Correspondence Project*, accessible by: <https://www.darwinproject.ac.uk/letter/DCP-LETT-13230.xml> [consultado pela última vez a 3 de Julho de 2017].
- \_\_\_\_; [1881b] *The Formation of Vegetable Mould through the Action of Worms, with Observations on their Habits*, John Murray, London.
- DAVIDSON, Donald; [1977] «The method of truth in metaphysics», P.A. French, T.E. Uehling & H. Wettstein (eds.), *Midwest Studies in Philosophy, II: Studies in Metaphysics*, University of Minnesota Press, Minneapolis, MN.
- \_\_\_\_; [1987] «Knowing One's Own Mind» (pp.441-458), *Proceedings and Addresses of the American Philosophical Association*, 60 (3).

- \_\_\_; [1999] «Intellectual Bibliography», in L. Hahn (ed.), *The Philosophy of Donald Davidson*, Library of Living Philosophers, Open Court, La Salle, IL.
- DAWKINS, Richard; [2003] «Introduction» (pp.ix-xxx), Charles Darwin, *The Origin of Species and The Voyage of the Beagle*, Everyman's Library & Alfred A. Knopf, Toronto.
- DAY, Rachel L., LALAND, Kevin, & ODLING-SMEE, F. John; [2003] «Rethinking Adaptation - the niche-construction perspective» (pp.80–95), *Perspectives in Biology and Medicine*, 46 (1).
- DEACON, Terrence; [2011] *Incomplete Nature: How Mind Emerged from Matter*, W. W. Norton & Company, Inc..
- DE CRUZ, Helen, BOUDRY, Maarten, DE SMEDT, Johan & BLANCHE, Stefaan; [2011] «Evolutionary Approaches to Epistemic Justification» (pp.517-535), *Dialectica*, 65 (4).
- DE CRUZ, Helen & DE SMEDT, Johan; [2012] «Evolved cognitive biases and the epistemic status of scientific beliefs» (pp. 411-429), *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 157 (3).
- DELBRÜCK, Max; [1986] *Mind from Matter? – An Essay on Evolutionary Epistemology*, Gunther S. Stent et al. (eds.), Blackwell Scientific Publications, Inc., Palo Alto, California.
- DENNETT, Daniel C.; [1987] *The Intentional Stance*, The MIT Press, Cambridge, MA.
- \_\_\_; [1991] *Consciousness Explained*, Little Brown & Co, Boston, MA.
- \_\_\_; [1995] *Darwin's Dangerous Idea: Evolution and the Meaning of Life*, Simon & Schuster.
- DE PRADO SALAS, Javier González; [2018] «Still Unsuccessful: The Unsolved Problems of Success Semantics» (pp.5-22), *Theoria*, 33 (1).
- DE WAAL, Frans; [2016] *Are we smart enough to know how smart animals are?*, W. W. Norton & Company, Inc..
- DIOGO, Rui; [2017] *Evolution Driven By Organismal Behavior: A Unifying View of Life, Function, Form, Mismatches, and Trends*, Springer International Publishing, Gewerbestrasse.
- DOBZHANSKY, Theodosius; [1973] «Nothing in Biology Makes Sense Except in the Light of Evolution» (pp.125–129), in *American Biology Teacher*, 35 (3).
- DOKIC, Jérôme & ENGEL, Pascal; [2002] *Frank Ramsey: Truth and Success*, Routledge.
- \_\_\_; [2005] «Ramsey's Principle Resituated», in Hallvard Lillehammer & D. H. Mellor (eds.) *Ramsey's Legacy*, Oxford University Press, Oxford.
- DWECK, Carol S.; [2009] «Why we don't need built-in misbeliefs» (pp.518-519), *Behavioral and Brain Sciences*, 32 (6).
- EDWARDS, Douglas; [2016] «Truth as a Relational Property» (pp.1-23), *Synthese*, 193 (11).
- EGBERT, Matthew D. et al. [2010] «A Minimal Model of Metabolism-Based Chemotaxis», *PLoS Computational Biology*, 6 (12).
- EINSTEIN, Albert; [1956] «Physics and Reality» (pp.349-82), *Journal of the Franklin Institute*, 221 (3).
- EL MOULDEN, Claire et al.; [2012] «What do humans maximize?» (pp.23-49), Samir Okasha & Ken Binmore (eds.), *Evolution and Rationality: Decisions, Co-operation and Strategic Behaviour*, Cambridge University Press, Cambridge.
- ENGEL, Pascal; [2004] «Truth and the aim of belief» (pp.77–97), Donald Gillies (ed.), *Laws and Models in Science*, King's College Publications, London.
- ENGELS, Friedrich [1876] «The Part Played by Labour in the Transition from Ape to Man», acedido pela última vez em 16/02/2018, <https://www.marxists.org/archive/marx/works/1876/part-played-labour/>.
- ENGLAND, Jeremy; [2013] «Statistical physics of self-replication» (pp.121923-1-8), *The Journal of Chemical Physics*, 139 (12).
- FALES, Evan; [1996] «Plantinga's case against naturalistic epistemology» (pp.432-451), *Philosophy of Science*, 63 (3).

- FELDMAN, Richard; [1988] «Rationality, Reliability, and natural selection» (pp.218-227), *Philosophy of Science*, 55 (2).
- FINE, Kit; [2001] «The Question of Realism» (pp.1-30), *Philosophers' Imprint*, 1.
- FITELSON, Branden & SOBER, Elliott; [1997] «Plantinga's Probability Arguments Against Evolutionary Naturalism» (pp.115-129), *Pacific Philosophical Quarterly*, 79.
- FLANAGAN, Owen; [2009] «“Can do” attitudes: Some positive illusions are not misbeliefs» (pp.519-520), *Brain and Behavioral Sciences*, 32 (6).
- FODOR, Jerry; [1981] «Three cheers for propositional attitudes», *Representations: Philosophical Essays on the Foundations of Cognitive Science*, The MIT Press.
- FODOR, Jerry & PIATTELLI-PALMARINI, Massimo; [2010] *What Darwin Got Wrong*, Farrar, Straus and Giroux.
- FORD DOOLITTLE, W.; [2014] «Natural selection through survival alone, and the possibility of Gaia» (pp.415-423), *Biology and Philosophy*, 29.
- FRISTON, Karl; [2010] «The free-energy principle: a unified brain theory?» (pp.127-138), *Nature Reviews Neuroscience*, 11, February.
- FRISTON, Karl J. & STEPHAN, Klaas E.; [2007] «Free-energy and the brain» (pp.417-458), *Synthese*, 159.
- GELL-MANN, Murray; [1994] *The Quark and the Jaguar*, Freeman & Co., New York.
- GENDLER, Tamar Szabó; [2008] «Alief and Belief» (pp.634-663), *Journal of Philosophy*, 105 (10).
- GIBSON, James J.; [1979] *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston.
- GIERE, Ronald N.; [1988] *Explaining Science – A Cognitive Approach*, The University of Chicago Press, Chicago.
- \_\_\_\_; [2006] «Modest Evolutionary Naturalism» (pp.52-60), *Biological Theory*, 1 (1).
- \_\_\_\_; [2010 (2006)] *Scientific Perspectivism*, The University of Chicago Press, Chicago & London.
- GIGERENZER, Gerd, G., Todd, P. M., & the ABC Group; [1999] *Simple heuristics that make us smart*, Oxford University Press, New York.
- GŁADZIEKEWSKI, Paweł; [2016] «Action guidance is not enough, representations need correspondence too: A plea for a two-factor theory of representation» (pp.13-25), *New Ideas in Psychology*, 40.
- GODFREY-SMITH, Peter; [1991] «Signal, Detection, and Action» (pp.709-722), *Journal of Philosophy*, 88 (12).
- \_\_\_\_; [1994] «A Continuum of Semantic Optimism» (pp.259-277), Stephen P. Stich & Ted A. Warfield (eds.), *Mental Representation: A Reader*, Blackwell.
- \_\_\_\_; [1998] *Complexity and the Function of Mind in Nature*, Cambridge University Press.
- \_\_\_\_; [2001a] «Organism, Environment and Dialectics» (pp.253-266), R. Singh, C. Krimbas, D. Paul, & J. Beatty (eds.), *Thinking About Evolution: Historical, Philosophical and Political Perspectives*, Volume 2 of a Festschrift for R. C. Lewontin, Cambridge University Press, Cambridge.
- \_\_\_\_; [2001b] «Three kinds of adaptationism» (pp. 335-357), S. H. Orzack, and E. Sober (eds.), *Adaptationism and Optimality*, Cambridge University Press, New York.
- \_\_\_\_; [2002] «Environmental Complexity and the Evolution of Cognition» (pp.233-249), Robert J. Sternberg & J. Kaufman (eds.), *The Evolution of Intelligence*, Lawrence Erlbaum.
- \_\_\_\_; [2009] *Darwinian Populations and Natural Selection*, Oxford University Press, Oxford.
- \_\_\_\_; [2014] «Adaptation, Construction, Function», *Philosophy of Biology*, Princeton University Press, Princeton, NJ.
- \_\_\_\_; [2016] *Other Minds - The Octopus, the Sea, and the Deep Origins of Consciousness*, Farrar, Straus & Giroux.
- GODFREY-SMITH, Peter & STERELNY, Kim; [2016] «Biological Information», *The Stanford Encyclopedia of Philosophy*, Summer 2016 Edition, Edward N. Zalta (ed.), URL = < <https://plato.stanford.edu/archives/sum2016/entries/information-biological/> >.

- GOLDMAN, Alvin & BEDDOR, Bob; [2015] «Reliabilist Epistemology», *The Stanford Encyclopedia of Philosophy*, Winter 2016 Edition, Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2016/entries/reliabilism/>>.
- GOODE, R. & GRIFFITHS, Paul E.; [1995] «The Misuse of Sober's Selection for/Selection of Distinction» (pp.99-108), *Biology and Philosophy*, 10 (1).
- GOULD, Stephen Jay & LEWONTIN, Richards; [1979] «The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme» (pp.581-598), *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 205 (1161), *The Evolution of Adaptation by Natural Selection*, September 21st.
- GOULD, S.J. & VRBA, E.S.; [1998] «Exaptation – A Missing Term in the Science of Form» (pp.52-71), David Hull & Michael Ruse (eds.), *The Philosophy of Biology*, Oxford University Press.
- GRIFFITHS, Paul E. & WILKINS, John S.; [2015] «Crossing the Milvian Bridge: When Do Evolutionary Explanations of Belief Debunk Belief?» (pp.201–231), P. Sloan, G. McKenny, & K. Eggelson (eds.), *Darwin in the 21st Century: Nature, Humanity, and God*, Notre Dame University Press, Notre Dame, IN.
- HACKING, Ian; [1983] *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*, Cambridge University Press, Cambridge.
- HARDIN, Garrett; [1968] «The Tragedy of the Commons» (pp.1243–1248), *Science*, 162 (3859).
- HASELTON, M. G. & NETTLE, D.; [2006] «The paranoid optimist: An integrative evolutionary model of cognitive biases» (p.47–66), *Personality and Social Psychology Review*, 10 (1).
- HOLLAND, John H.; [2006], «Studying Complex Adaptive Systems» (pp.1–8), *Journal of Systems Science and Complexity*, 19 (1).
- HOOKE, Clifford Alan; [1975] «The Information-Processing Approach to the Brain-Mind and its Philosophical Ramifications» (pp.1-15), *Philosophy and Phenomenological Research*, 36 (1).
- \_\_\_\_; [1995] *Reason, Regulation, and Realism -Towards a Regulatory Systems Theory of Reason and Evolutionary Epistemology*, SUNY Press, Albany, NY.
- \_\_\_\_; [2009] «Interaction and bio-cognitive order» (pp. 513–546), *Synthese*, 166.
- \_\_\_\_; [2011] «Introduction to Philosophy of Complex Systems: A: Part A: Towards a framework for complex systems» (pp.3-90), Clifford A. Hooker (ed.), *Philosophy of Complex Systems*, North Holland, of Elsevier B.V., Waltham, MA.
- \_\_\_\_; [2015] «Forward», *Moreno & Mossio's 2015*.
- HUMPHREY, Nicholas K.; [1976] «The social function of intellect» (pp.303- 317), P. P. G. Bateson & R. A. Hinde (eds.), in *Growing Points in Ethology*, Cambridge University Press, Cambridge.
- \_\_\_\_; [2004] «Placebo effect» (pp.735-6), Richard L. Gregory (ed.), *The Oxford Companion to the Mind*, Oxford University Press, New York.
- HUXLEY, Julian; [1942] *Evolution: The Modern Synthesis*, Allen & Unwin, London.
- HUXLEY, Thomas H.; [1870] «On the educational value of the natural history sciences», *Lay sermons, addresses, and reviews*, Macmillan, London.
- ISLER, Karin & VAN SCHAİK, Carel P.; [2006] «Metabolic costs of brain size evolution» (pp.557–560), *Biology Letters*, 2 (4).
- JABLONSKI, David; [2008] «Species Selection: Theory and Data» (pp.501-524), *Annual Review of Ecology, Evolution, and Systematics*, 39.
- JACOB, François; [1982] *The Possible and the Actual*, University of Washington Press, Seattle.
- JACOB, Pierre; [2014] «Intentionality», *The Stanford Encyclopedia of Philosophy*, Winter 2014 Edition, Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2014/entries/intentionality/>>.
- JÉKELY, G., KEIJZER, F. & GODFREY-SMITH, Peter; [2015] «An Option Space for Early Neural Evolution» (pp.1-10), *Philosophical Transactions of the Royal Society B*, 370.
- KAHANE, Guy; [2011] «Evolutionary Debunking Arguments» (pp.103-125), *Noûs*, 45 (1).

- KANEKO, Kuniyiko; [2012] «Evolution of Robustness and Plasticity under Environmental Fluctuation: Formulation in terms of Phenotypic Variances» (pp.687–705), *Journal of Statistical Physics*, 148 (4).
- KANT, Immanuel; [1951 (1790)] «§65 Analytic of the Teleological Judgment», *Critique of Judgment*, J. H. Bernard (ed.), Hafner Press, New York, NY.
- KAUFFMAN, Stuart; [1993] *The Origins of Order*, Oxford University Press, Oxford.  
 \_\_\_\_; [2000] *Investigations*, Oxford University Press, Oxford.
- KIMURA, Motoo; [1983] *The Neutral Theory of Molecular Evolution*, Cambridge University Press, Cambridge.
- KITCHER, Philip; [1985] *Vaulting Ambition – Sociobiology and the Quest for Human Nature*, The MIT Press.  
 \_\_\_\_; [1990] «The Division of Cognitive Labor» (pp.5-22), *The Journal of Philosophy*, 87 (1).  
 \_\_\_\_; [1993] «Function and design» (pp.379-397), *Midwest Studies in Philosophy*, 18.
- KORNBLITH, Hilary; [1993] *Inductive Inference and Its Natural Ground - An Essay in Naturalistic Epistemology*, The MIT Press.  
 [1994] «Naturalism: Both Metaphysical and Epistemological» (pp.39-52), *Midwest Studies in Philosophy*, 19.  
 \_\_\_\_; [2002] *Knowledge and its Place in Nature*, Oxford University Press.
- KUHN, Thomas; [2012 (1962)] *The Structure of Scientific Revolutions*, The University of Chicago Press, Chicago.
- KVANVIG, Jonathan; [2003] *The Value of Knowledge and the Pursuit of Understanding*, Cambridge University Press, Cambridge.  
 \_\_\_\_; [2005] «Truth is not the primary epistemic goal» (pp.285-296), in Matthias Steup & Ernest Sosa (eds.), *Contemporary Debates in Epistemology*, Blackwell, Malden, MA.
- LADYMAN, James & ROSS, Don, with SPURRETT, David & COLLIER, John; [2007] *Every Thing Must Go*, Oxford University Press, Oxford.
- LALAND, Kevin N. & BROWN, Gillian R.; [2011] *Sense & Nonsense – Evolutionary perspectives on human behaviour*, Oxford University Press, Oxford.
- LALAND, Kevin A. et al.; [2014] «Does Evolutionary Theory Need a Rethink?» (pp.161–164), *Nature*, 514 (7521).
- LALAND, Kevin A. et al.; [2015] «The extended evolutionary synthesis: its structure, assumptions and predictions» (pp.1-14), *The Royal Society Publishing Proceedings B.*, 282 (1813).
- LAUDAN, Larry; [1981] «A Confutation of Convergent Realism» (pp.19-49), *Philosophy of Science*, 48 (1).
- LAW, Richard; [1979] «Optimal Life Histories Under Age-Specific Predation» (pp.399-417), *The American Naturalist*, 114 (3).
- LETTVIN, J. Y., MATURANA, H.R., MCCULLOCH, W.S. & PITTS, W.H.; [1959] «What the frog's eye tells the frog's brain», *Proceedings of the IRE*, 47 (11).
- LEVINS, Richard; [1966] «The Strategy of Model-Building in Population Biology» (pp.421-31), *American Scientist*, 54.
- LEVINS, Richard & LEWONTIN, Richard; [1983] *The Dialectical Biologist*, Harvard University Press, Cambridge, MA.
- LEWIS, David; [1970] «How to define theoretical terms» (pp.426–446), *Journal of Philosophy*, 67.
- LEWONTIN, Richard; [1978] «Adaptation» (pp.212-8), *Scientific American*, 239 (3).  
 \_\_\_\_; [1982], «Organism and Environment» (pp.151-170), in H.C. Plotkin (ed.) *Learning, Development, and Culture: Essays in Evolutionary Epistemology*, Wiley, Chichester, England.  
 \_\_\_\_; [1983] «The organism as subject and object of evolution», Levins & Lewontin [1983].  
 \_\_\_\_; [1997] «Billions and Billions of Demons» (pp.28-32), *The New York Review of Books*, 44 (1).

\_\_\_\_; [1998] «The evolution of cognition: Questions we will never answer» (pp.107-132), in D. Scarborough and S. Sternberg (eds), *An invitation to cognitive science, Volume 4: Methods, models, and conceptual issues*, The MIT Press, Cambridge, MA.

LIEBENBERG, Louis; [2013] «Tracking Science – The Origin of Scientific Thinking in our Paleolithic Ancestors» (pp.18-23), *Skeptical Magazine*, 18 (3).

LIPTON, Peter; [2000] «Tracking Track Records» (pp.179–205), *Aristotelian Society Supplementary Volume*, 74 (1).

\_\_\_\_; [2004 (1991)] *Inference to the best explanation*, Routledge.

LLOYD, Elisabeth; [2017] «Units and Levels of Selection», *The Stanford Encyclopedia of Philosophy*, Summer 2017 Edition, Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2017/entries/selection-units/>>.

LOEWER, Barry; [1997] «A Guide to Naturalizing Semantics» (pp.108-126), B. Hale and C. Wright (eds.), *A Companion to the Philosophy of Language*, Blackwell, Malden, MA.

LORENZ, Konrad; [1975 (1941)] «Kant's doctrine of the a priori in light of contemporary biology», Richard J. Evans (trad. & ed.) *Konrad Lorenz, The Man and his Ideas*, Harcourt Brace Jovanovich, New York.

\_\_\_\_; [1973] *Behind the Mirror: A Search for a Natural History of Human Knowledge*, Piper & Co. Verlag.

LOTKA, Alfred J.; [1922a] «Contribution to the energetic of evolution» (pp.147–151), *Proceedings of the National Academy of Sciences*, 8 (6).

\_\_\_\_; [1922b] «Natural Selection as a Physical Principle» (pp.151-154), *Proceedings of the National Academy of Sciences of the United States of America*, 8 (6).

MADDY, Penelope; [2007] *Second Philosophy: A Naturalistic Method*, Oxford University Press.

MAKI, Nazli *et al.*; [2000] «Motility and Chemotaxis of Filamentous Cells of *Escherichia coli*» (p.4337–4342), *Journal of Bacteriology*, 182 (15).

MATURANA, Humberto R. & VARELA, Francisco; [1972] *Autopoiesis and Cognition – The Realization of the Living*, Boston Studies in the Philosophy of Science, D. Reidel Publishing Company, Boston, MA.

MAXWELL, Grover; [1962] «The Ontological Status of Theoretical Entities» (pp.3–15), in H. Feigl and G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science III*, University of Minnesota Press, Minneapolis.

MAYNARD SMITH, John; [1975] *The Theory of Evolution*, Penguin Books, Harmondsworth, England.

\_\_\_\_; [1978] «Optimization theory in evolution» (pp.31-56), *Annual Review of Ecology and Systematics*, 9.

MCKAY, Richard & DENNETT, Daniel C.; [2009] «Evolution of Misbelief» (pp.493-510), *Behavioral and Brain Sciences*, 32 (6).

MCKITTRICK, Jennifer; [2003] «A Case for Extrinsic Dispositions» (pp.155–174), *Australasian Journal of Philosophy*, 81.

METHVEN, Stephen J.; [2015] *Frank Ramsey and the Realistic Spirit*, Palgrave Macmillan, Basingstoke.

MESOUDI, Alex; [2016a] «Cultural Evolution: A Review of Theory, Findings and Controversies» (pp.481–497), *Evolutionary Biology*, 43 (4).

\_\_\_\_; [2016b] «Cultural evolution - integrating psychology, evolution and culture» (pp.17–22), *Current Opinion in Psychology*, 7.

\_\_\_\_; [2017] «Pursuing Darwin's curious parallel: Prospects for a science of cultural evolution» (pp.7853–7860), *PNAS*, 114 (30).

MESOUDI, Alex *et al.*; [2013] «The Cultural Evolution of Technology and Science» (pp.193-216), in Peter J. Richerson & Morton H. Christiansen (eds.), *Cultural Evolution: Society, Technology, Language, and Religion*, Strüngmann Forum Reports, The MIT Press, Cambridge, MA.

MIŁKOWSKI, Marcin; [2015] «Function and Causal Relevance of Content», in *New Ideas in Psychology*

MILLIKAN, Ruth G.; [1984] *Language, Thought, and other Biological Categories – New Foundations for Realism*, The MIT Press, Cambridge, MA.



- \_\_\_\_; [1989] «Biosemantics» (pp.281-97), in *Journal of Philosophy*, 86.
- \_\_\_\_; [1993] *White queen psychology and other essays for Alice*, The MIT Press.
- MITCHELL, Sandra; [2003] *Biological Complexity and Integrative Pluralism*, Cambridge University Press, Cambridge.
- \_\_\_\_; [2004] «Why Integrative Pluralism?», *Emergence: Complexity and Organization*, 1.
- MITHEN, Steven; [2004 (2002)] «Human evolution and the cognitive basis of science» (pp.23-40), Peter Carruthers, Stephen Stich, & Michael Siegal (eds.), *The Cognitive Basis of Science*, Cambridge University Press, Cambridge.
- MORENO, Alvaro & MOSSÍO, Matteo; [2015] *Biological Autonomy – A Philosophical and Theoretical Enquiry*, Springer, Dordrecht.
- NANAY, Bence; Nanay [2011] «Replication without replicators» (pp.455–477), *Synthese*, 179 (3).
- \_\_\_\_; [2013] «Success semantics: the sequel» (pp.151-65), *Philosophical Studies*, 165 (1).
- NEANDER, Karen; [1991] «The teleological notion of “function”» (pp.454-468), *Australasian Journal of Philosophy*, 69 (4).
- \_\_\_\_; [1995] «Misrepresenting & Malfunctioning» (pp.109-141), *Philosophical Studies*, 79 (2).
- NESSE, Randolph M.; [2005] «Natural selection and the regulation of defensive responses» (pp.88-105), *Evolution and Human Behavior*, 26.
- NIETZSCHE, Friedrich; [1974 (1887)] *The Gay Science – with a prelude in rhymes and an appendix of songs*, Vintage Books, New York.
- NIINILUOTO, Ilkka; [1999] *Critical Scientific Realism*, Clarendon Library of Logic and Philosophy, Oxford University Press, Oxford.
- NOWAK, Martin; [2006] «Five Rules for the Evolution of Cooperation» (pp.1560–1563), *Science*, 314 (5805).
- NOWAK, Martin A., TARNITA, Corina E. & WILSON, Edward O.; [2010] «The evolution of eusociality» (pp.1057–1062), *Nature*, 466.
- NOZICK, Robert; [1993] *The Nature of Rationality*, Princeton University Press, Princeton, NJ.
- ODLING-SMEE, F. J. *et al.*; [2003] «Rethinking Adaptation - the niche-construction perspective» (pp.80–95), *Perspectives in Biology and Medicine*, 46 (1).
- OHTA, Tomoko; [1973] «Slightly Deleterious Mutant Substitutions in Evolution» (pp.96–98), *Nature*, 246 (5428), November 9<sup>th</sup>.
- OKASHA, Samir; [2006] *Evolution and the Levels of Selection*, Oxford University Press, Oxford.
- ORILIA, Francesco & SWOYER, Chris; [2016] «Properties», *The Stanford Encyclopedia of Philosophy*, Winter 2016 Edition, Edward N. Zalta (ed.), URL = < <https://plato.stanford.edu/archives/win2016/entries/properties/> >.
- ORZACK, Steven & SOBER, Elliott; [1994] «Optimality models and the test of Adaptationism» (pp.361-80), *The American Naturalist*, 143 (3).
- \_\_\_\_; [1996] «How to Formulate and Test Adaptationism» (pp.202-210), *The American Naturalist*, 148 (1).
- PAPINEAU, David; [1987] *Reality and Representation*, Blackwell, Oxford.
- \_\_\_\_; [1990] «Truth and Teleology» (pp.21-43), D. Knowles (ed.), *Explanation and its Limits*, Cambridge University Press.
- \_\_\_\_; [1993] *Philosophical Naturalism*, Blackwell, Oxford.
- \_\_\_\_; [2001] «The status of teleosemantics, or how to stop worrying about swampman» (pp.279-89), *Australasian Journal of Philosophy*, 79 (2).
- \_\_\_\_; [2003] «Is representation rife?» (pp.107-123), *Ratio – An International Journal of Analytic Philosophy*, 16 (2).
- \_\_\_\_; [2016] «Teleosemantics» (pp.95-120), David Livingstone Smith (ed.), *How Biology Shapes Philosophy: New Foundations for Realism*, Cambridge University Press, Cambridge.

- PAPPAS, George; [2017] «Internalist vs. Externalist Conceptions of Epistemic Justification», in *The Stanford Encyclopedia of Philosophy*, Fall 2017 Edition, Edward N. Zalta (ed.), URL = < <https://plato.stanford.edu/archives/fall2017/entries/justep-intext/> >.
- PASCAL, Robert & PROSS, Addy; [2016] «The Logic of Life» (pp.507–513), *Origins of Life and Evolution of Biospheres*, 46.
- PEACOCK, Kent A.; [2011] «The Three Faces of Ecological Fitness» (pp.99-105), *Studies in History and Philosophy of Biological and Biomedical Sciences*, 42 (1).
- PEIRCE, Charles Sanders; [1903] «The Fixation of Belief», *Collected Papers of Charles Sanders Peirce*, vol. 5 *Pragmatism and Pragmaticism* (1934), Charles Hartshorne & Paul Weiss (eds.), Harvard University Press, Cambridge, MA.
- \_\_\_\_; [1955 (1883)] «The General Theory of Probable Inference» (pp.190-217), Justus Buchler (ed.), *Philosophical Writings of Peirce*, Dover.
- PERRY, John; [1993] *The Problem of the Essential Indexical and Other Essays*, Oxford University Press, Oxford.
- PLANTINGA, Alvin; [1993] *Warrant and Proper Function*, Oxford University Press, Oxford.
- PLOTKIN, Henry C.; [1987] «Evolutionary epistemology as a science» (pp.295–313), *Biology and Philosophy*, 2 (3).
- PINKER, Steven; [2005] «So how does the mind work?» (pp.1-24), *Mind and Language*, 20 (1).
- \_\_\_\_; [2009] *How the Mind Works*, W. W. Norton & Company, Inc.
- PINKER, Steven et al.; [1997] «Evolutionary Psychology: An Exchange», *The New York Review of Books*, <http://www.nybooks.com/articles/1997/10/09/evolutionary-psychology-an-exchange/> [consultado pela última vez a 5/02/2018].
- POPPER, Karl; [1976] *Unended Quest. An Intellectual Autobiography*, Open Court, LaSalle, IL.
- \_\_\_\_; [1978] «Natural Selection and the Emergence of Mind» (pp.339-355), *Dialectica*, 32.
- PRICE, Huw; [1989] «Defending Desire-as-Belief» (pp.119-27), *Mind*, 98.
- PRIGOGINE, Ilya; [1980] *From Being to Becoming: Time and complexity in the Physical Sciences*, W.H. Freeman & Company.
- PRIGOGINE & STENGERS, Isabelle; [1984] *Order out of Chaos: The Evolutionary Paradigm and the Physical Sciences*, Bantam.
- PSILLOS, Stathis; [1999] *Scientific Realism – How Science Tracks Truth*, Routledge.
- PUTNAM, Hilary; [1979 (1975)] «What is mathematical truth?» (pp.60-79), in Hilary Putnam (ed.), *Philosophical Papers, Vol.1: Mathematics, Matter and Method*.
- \_\_\_\_; [1981] *Reason, Truth and History*, Cambridge University Press, Cambridge.
- QUINE, Willard van Orman; [1969] *Ontological Relativity and Other Essays*, Columbia University Press.
- \_\_\_\_; [1975] «The Nature of Natural Knowledge» (pp.1975–67), Samuel D. Guttenplan (ed.), *Mind and Language*, Clarendon Press.
- \_\_\_\_; [1986] «Reply to Morton White» (pp.663-5), L. Hahn, P. Schilpp (eds.), *The Philosophy of W.V.O. Quine*, The Library of Living Philosophers, Open Court, La Salle, IL.
- RAMSEY, Frank P.; [1927] «Facts and Propositions» (pp.153-170), *Proceedings of the Aristotelian Society*, 7 (1).
- \_\_\_\_; [1929] «Truth and Probability», Richard B. Braithwaite (ed.), *The Foundations of Mathematics and other Logical Essays*, Routledge & Kegan Paul Ltd, London.
- \_\_\_\_; [1931] «Knowledge» (pp.126–128), R.B. Braithwaite (ed.), *Foundations of Mathematics and Other Logical Essays*, Kegan Paul, London.
- RESCHER, Nicholas; [1989] *A Useful Inheritance: Evolutionary Aspects of the Theory of Knowledge*, Rowman & Littlefield Publishers.

- RICHERSON, Peter J. & BOYD, Robert; [2000] «Climate, Culture, and the Evolution of Cognition» (pp. 329—46), Celia Heyes & Ludwig Huber (eds.), *The Evolution of Cognition*, The Vienna Series in Theoretical Biology, The MIT Press.
- RIEDL, Rupert; [1984] *Biology of Knowledge – The Evolutionary Basis of Reason*, A Wiley-Interscience Publication, transl. from the 3<sup>rd</sup> German Edition by Paul Foulkes, John Wiley & Sons, Chichester and New York.
- ROSEN, Robert; [1970] *Dynamical Systems Theory in Biology: Stability Theory and Its Applications*, Wiley-Interscience.
- \_\_\_\_; [1985] *Anticipatory Systems: Philosophical, Mathematical, and Methodological Foundations*, Pergamon Press.
- ROSENBERG, Alexander & BOUCHARD, Frédéric; [2015] «Fitness», *The Stanford Encyclopedia of Philosophy*, Fall 2015 Edition, Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2015/entries/fitness/>>.
- ROSSLENBROICH, Bernd; [2014] *On the Origin of Autonomy - A New Look at the Major Transitions in Evolution*, Philosophy and Theory of the Life Sciences, Volume 5, Springer, New York, NY.
- RUPERT, Robert D.; [2009] *Cognitive Systems and the Extended Mind*, Oxford University Press, Oxford.
- RUSE, Michael; [1986] *Taking Darwin Seriously*, Prometheus Books.
- RUSSELL, Bertrand; [1906-7] «The Monistic Theory of Truth» (pp.150-169), part of «The Nature of Truth», *Proceedings of the Aristotelian Society*, disponível em <http://bertrandrussellsocietylibrary.org/br-pe/br-pe-ch6.html> [consultado pela última vez a 24/03/2017].
- \_\_\_\_; [1959] *The Problems of Philosophy*, Oxford University Press, Oxford.
- SAGE, James; [2004] «Truth-Reliability and the Evolution of Human Cognitive Faculties» (pp. 95-106), *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, Vol. 117, No. 1/2, Selected Papers from the American Philosophical Association, Pacific Division, 2003 January Meeting.
- SAHLIN, Nils-Eric; [1990] *The Philosophy of Frank Ramsey*, Cambridge University Press, Cambridge.
- SANKEY, Howard; [2014] «Scientific Realism and Basic Common Sense» (p.11-24), *Kairos*, 10.
- SANTOS, Gil. C.; [2015a] «Ontological Emergence: How is That Possible? Towards a new Relational Ontology», *Foundations of Science*.
- \_\_\_\_; [2015b] Upward and Downward Causation from a Relational-Horizontal Ontological Perspective» (pp.23-40), *Axiomathes*, 25 (1).
- SCHAFFER, Jonathan; [2010] «The Internal Relatedness of All Things» (pp.341-376), *Mind*, 119 (474).
- SCHRÖDINGER, Erwin; [1944] *What Is Life? The Physical Aspect of the Living Cell*, Cambridge University Press.
- SCOTT, R. M. & BAILLARGEON, R.; [2009] «Which penguin is this? - Attributing false beliefs about object identity at 18 months» (pp.1172-96), *Child Development*, 80 (4).
- SHAPIRO, Lawrence A.; [1997] «The nature of nature: rethinking naturalistic theories of intentionality» (p.309-22), in *Philosophical Psychology*, 10 (3).
- SHEA, Nicholas; [2007] «Representation in the genome and in other inheritance systems» (pp.313–31), *Biology & Philosophy*, 22 (3).
- \_\_\_\_; [2012] «Reward Prediction Error Signals are Meta-Representational» (pp.314-41), *Noûs*, 48 (2).
- SIMPSON, George Gaylord; [1963] *This view of life: the world of an evolutionist*, Harcourt, Brace & World, New York, NY.
- SKLAR, Lawrence; [2010] «I'd Love to Be a Naturalist—if Only I Knew What Naturalism Was» (pp.1121-1137), *Philosophy of Science*, 77 (5).
- SMOLIN, Lee; [1997] *The Life of the Cosmos*, Oxford University Press, Oxford.
- SOBER, Elliott; [1980] «Evolution, population thinking, and essentialism» (pp.350-83), *Philosophy of Science*, 47 (3).

- \_\_\_\_; [1981] «The Evolution of Rationality» (pp.95-120), *Synthese*, 46.
- \_\_\_\_; [1993a (1984)] *The Nature of Selection*, The University of Chicago Press, Chicago.
- \_\_\_\_; [1993b] *Philosophy of Biology*, Oxford University Press, Oxford.
- \_\_\_\_; [1994] *From a biological point of view – essays in evolutionary philosophy*, Cambridge University Press, Cambridge.
- \_\_\_\_; [1996] «Evolution and Optimality: Feathers, Bowling Balls, and the Thesis of Adaptationism» (pp.41-57), *Philosophic Exchange*, 26 (1):2.
- \_\_\_\_; [1997] «Is the Mind an Adaptation for Coping with Environmental Complexity?» (pp.539–550), *Biology and Philosophy*, 12.
- \_\_\_\_; [2001] «The Two Faces of Fitness» (pp.309-321), R. Singh, D. Paul, C. Krimbas, & J. Beatty (eds.), *Thinking about Evolution: Historical, Philosophical, and Political Perspectives*, Cambridge University Press, Cambridge.
- SOUTHGATE, V., SENJU, A., & CSIBRA, G.; [2007] «Action Anticipation through Attribution of False Belief by Two-Year-Olds» (pp.587-92), *Psychological Science*, 18(7).
- SPENCER, Herbert; [1852] «A theory of population, deduced from the General Law of Animal Fertility», [http://www.victorianweb.org/science/science\\_texts/spencer2.html](http://www.victorianweb.org/science/science_texts/spencer2.html) [consultado pela última vez a 15/09/2017].
- STALEY, Kent W.; [2014] *An Introduction to the Philosophy of Science*, The Cambridge University Press, Cambridge.
- STALNAKER, Robert; [1987] *Inquiry*, Bradford Books, The MIT Press, Cambridge, MA.
- STEFFEN, Will *et al.*; [2011] «The Anthropocene: conceptual and historical perspectives» (pp.842–867), *Philosophical Transactions of the Royal Society A*, 369.
- STEPHENS, Christopher L.; [2001] «When is it selectively advantageous to have true beliefs? Sandwiching the better safe than sorry argument» (pp.161–189), *Philosophical Studies*, 105 (2).
- STERELNY, Kim; [2003] *Thought in a Hostile World – The Evolution of Human Cognition*, Wiley-Blackwell.
- \_\_\_\_; [2012] «From fitness to utility» (pp.246-273), Samir Okasha & Ken Binmore (eds.), *Evolution and Rationality: Decisions, Co-operation, and Strategic Behaviour*, Cambridge University Press, Cambridge.
- STERELNY, Kim & FRASER, Ben; [2016] «Evolution and Moral Realism» (pp. 981-1006), *The British Journal for the Philosophy of Science*, 68 (4).
- STEWART, J.; [1996] «Cognition = Life: Implications for higher-level cognition» (pp.311-326), *Behavioural Processes*, 35.
- STICH, Stephen; [1990] *The Fragmentation of Reason*, The MIT Press, Cambridge, MA.
- STOUT, Dietrich; [2010] «The Evolution of Cognitive Control» (pp.614-630), *Topics in Cognitive Science*, 2 (4).
- TANG, Weng Hong; [2014] «Success Semantics and Partial Belief» (pp.17-22), *Journal of Philosophical Research*, 39.
- TAYLOR, Shelley & BROWN, Jonathon; [1988] «Illusion and Well-Being - A Social Psychological Perspective on Mental Health» (pp.193-210), in *Psychological Bulletin*, 103 (2).
- TAYLOR, Shelly E.; [1989] *Positive illusions: Creative self-deception and the healthy mind*, Basic Books, New York, NY.
- TĚMKIN, Ylia & ELDREDGE, Niles; [2015] «Networks and Hierarchies: Approaching Complexity in Evolutionary Theory» (pp.183-226), E. Serrelli & N. Gontier (eds.), *Macroevolution: Explanation, interpretation and evidence*, Springer.
- THOMPSON, Francis; [1917] «The Mystress of Vision», Nicholson & Lee (eds.), *The Oxford Book of English Mystical Verse*, Oxford.
- THOMSON, Paul; [1995] «Evolutionary Epistemology and Scientific Realism» (pp.165-191), *Journal of Social and Evolutionary Systems*, 18 (2).

- TOMAN, & FLEGR, ; [2017] «Stability-based sorting: The forgotten process behind (not only) biological evolution» (pp.29-41), *Journal of theoretical Biology*, 435.
- TRIMMER, P.C. & HOUSTON, A.I.; [2014] «An evolutionary perspective on information processing» (pp.312–330), *Topics in Cognitive Science*, 6 (2).
- TURING, Alan; «The Chemical Basis of Morphogenesis» (pp.37-72), *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 237 (641).
- VAN FRAASSEN, Bas; [1980] *The Scientific Image*, Oxford University Press, Oxford.
- VAN GELDER, Tim; [2005] «Enhancing and augmenting human reasoning», António Zilhão (ed.), *Evolution, Rationality and Cognition – A cognitive science for the twenty-first century*, Routledge.
- VAN VALEN, Leigh; [1975] «Life death, and energy of a tree» (pp. 259–269), *Biotropica*, 7.  
 \_\_\_\_; [1989] «Three Paradigms of Evolution» (pp.1-17), *Evolutionary Theory*, 9.
- VARELA, Francisco J.; [1979] *Principles of biological autonomy*, North Holland, New York, NY.
- VICKERS, Peter; [2016] «Understanding the selective realist defence against the PMI» (pp.1-12), *Synthese, New Thinking About Scientific Realism*.
- VLERICK, Michael; [2017] «How our biology constrains our Science» (pp.31-53), *Kairos. Journal of Philosophy & Science*, 18 (1).
- VLERICK, Michael & BROADBENT, Alex; [2015] «Evolution and Epistemic Justification» (pp.185-203), *Dialectica*, 69.
- VOLLMER, Gerhard; [1984] «Mesocosm and Objective Knowledge - On Problems Solved by Evolutionary Epistemology» (pp.69-121), in Franz Wuketits (ed.), *Concepts and Approaches in Evolutionary Epistemology, Theory and Decision Library (An International Series in the Philosophy and Methodology of the Social and Behavioral Sciences)*, vol 36. Springer, Dordrecht.  
 \_\_\_\_; [2005 (1975)] *Teoría Evolucionista del Conocimiento*, Editorial Comares, Granada.
- VON BERTALANFFY; Ludwig; [1952] *Problems of life: an evaluation of modern biological and scientific thought*, Watts.  
 [1968] *General Systems Theory: Foundations, Development, Application*, George Braziller, Inc., New York, N.Y..
- VON UEXKÜLL, Jakob; [1934] «A Stroll through the World of Animals and Men – A Picture Book of Invisible Worlds» (pp.5-80), Claire H. Schiller (ed.), *Instinctive Behavior – The Development of a Modern Concept*, International Universities Press, Inc., New York.  
 \_\_\_\_; [1982] «The Theory of Meaning» (pp.25-82), *Semiotica* 42 (1).
- VOLTAIRE, pseud. of François-Marie Arouet; [1986 (1759)] *Candide ou l'optimisme: conte philosophique*, Claude Blum (ed.), Hachette, Paris.
- VRBA, Elisabeth S. & GOULD, Stephen Jay; [1986] «The hierarchical expansion of sorting and selection: sorting and selection cannot be equated» (pp.217-228), *Paleobiology*, 12 (2).
- WADDINGTON, C. H.; [1968] «The basic ideas in Biology» (pp.238-253), *Biological Theory*, 3 (3).
- WADE, Michael; [2005] «Evolutionary Genetics», in *The Stanford Encyclopedia of Philosophy*, Fall 2008 Edition, Edward N. Zalta (ed.), URL = < <https://plato.stanford.edu/archives/fall2008/entries/evolutionary-genetics/>>.
- WAGNER, Andreas; [2005] «Robustness, Evolvability, and Neutrality» (pp.1772–8), *FEBS Letters*, 579 (8).
- WALLACE, Alfred Russel; [1866] «Letter no. 5140», *Darwin Correspondence Project*, <http://www.darwinproject.ac.uk/DCP-LETT-5140> [consultado pela última vez a 10/11/2017].
- WEGNER, Daniel; [2002] *The Illusion of Conscious Will*, The MIT Press, Cambridge, MA.
- WHYTE, Jamie T.; [1990] «Success Semantics» (pp. 149-157), *Analysis*, 50 (3).  
 \_\_\_\_; [1993] «Purpose and Content» (pp.45-60), *British Journal of Philosophy of Science*, 44.
- WILLIAMS, Neil E.; [2010] «Puzzling powers: the problem of fit» (pp.84–105), Anna Marmodoro (ed.), *The Metaphysics of Powers: Their Grounding and Their Manifestations*, Routledge.

- WILSON, Edward O.; [1971] *The Insect Societies*, Belknap Press, Cambridge, MA.
- \_\_\_; [1974] «The Perfect Societies» (pp.54-5), *Science*, 184 (4132).
- \_\_\_; [1998] *Consilience: The Unity of Knowledge*, Vintage Books, New York.
- WIMSATT, William; [1997] «Functional organization, functional analogy, and functional inference» (pp.2-32), *Evolution and Cognition*, 3 (2).
- \_\_\_; [2007a] «On Building Reliable Pictures with Unreliable Data: an Evolutionary and Developmental Coda for the New Systems Biology», in *Systems Biology: Philosophical Foundations*, F.C. Boogerd, F.J. Bruggeman, J.-H.S. Hofmeyer, and H.V. Westerhoff (eds.), Reed-Elsevier, Amsterdam.
- \_\_\_; [2007b] *Re-engineering Philosophy for Limited Beings – Piecewise Approximations to Reality*, Harvard University Press, Harvard.
- WINNING, Jason & BECHTEL, William; [2016] «Review of *Biological Autonomy*» (pp.446-452), *Philosophy of Science*, 83 (3).
- WOLPERT, Daniel; [2011] «The real reason for brains», *Ted Talks*, [https://www.ted.com/talks/daniel\\_wolpert\\_the\\_real\\_reason\\_for\\_brains?language=pt](https://www.ted.com/talks/daniel_wolpert_the_real_reason_for_brains?language=pt) [consultado pela última vez a 18/11/2017].
- WORRALL, John; [1999] «Two cheers for naturalised philosophy of science – or: Why Naturalised Philosophy of Science Is Not the Cat's Whiskers» (pp.339-61), *Science and Education*, 8 (4).
- WRENN, Chase B.; [2011] «Practical success and the nature of truth» (pp.451–470), *Synthese*, 181.
- WRIGHT, Henry W.; [1913] «Practical Success as the Criterion of Truth» (pp.606-622), *The Philosophical Review*, 22 (6).
- WUKETITS, Franz; [1990] *Evolutionary Epistemology and Its Implications for Humankind*, SUNY Series in Philosophy and Biology, State University of New York Press.
- YEATS, William Butler; [1939] *Letter to Lady Elizabeth Pelham*, 4/January/1939.
- ZILHÃO, António; [2010] *Pensar com Risco – 25 Lições de Lógica Indutiva*, Imprensa Nacional – Casa da Moeda, Lisboa.